



Isotropic Multiple Scattering Processes on Hyperspheres

Nicolas Le Bihan, Florent Chatelain, Jonathan Manton

► To cite this version:

Nicolas Le Bihan, Florent Chatelain, Jonathan Manton. Isotropic Multiple Scattering Processes on Hyperspheres. IEEE Transactions on Information Theory, 2016, 62 (10), pp.5740-5752. 10.1109/TIT.2015.2508932 . hal-01322340

HAL Id: hal-01322340

<https://hal.science/hal-01322340>

Submitted on 27 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Isotropic Multiple Scattering Processes on Hyperspheres

Nicolas Le Bihan, Florent Chatelain, and Jonathan H. Manton, *IEEE fellow*

Abstract

This paper presents several results about isotropic random walks and multiple scattering processes on hyperspheres \mathbb{S}^{p-1} . It allows one to derive the Fourier expansions on \mathbb{S}^{p-1} of these processes. A result of unimodality for the multiconvolution of symmetrical probability density functions (pdf) on \mathbb{S}^{p-1} is also introduced. Such processes are then studied in the case where the scattering distribution is von Mises Fisher (vMF). Asymptotic distributions for the multiconvolution of vMFs on \mathbb{S}^{p-1} are obtained. Both Fourier expansion and asymptotic approximation allows us to compute estimation bounds for the parameters of Compound Cox Processes (CCP) on \mathbb{S}^{p-1} .

Index Terms

Isotropic random walk on \mathbb{S}^{p-1} , Compound Cox Processes on \mathbb{S}^{p-1} , von Mises-Fisher distribution, Fourier series expansion on hyperspheres, multiple scattering, Cramer-Rao lower bounds.

I. INTRODUCTION

Mixtures of von Mises-Fisher (vMF) distributions are models used in applications ranging from MRI data analysis [1] to radiation therapy beam direction clustering [2] and speaker clustering [3]. The finite mixture case was originally studied in [4] for data clustering on hyperspheres. All the above mentioned contributions made use of EM algorithms for the estimation of mixtures weights, vMF distribution parameters or number of mixture component.

In this paper, we consider particular countably infinite mixtures of directional distributions where the components of the mixture are multiply convolved distribution of unit vector in \mathbb{R}^p , *i.e.* elements of \mathbb{S}^{p-1} , and where the weights are controlled by a Cox process.

The proposed approach is valid for any dimension p , even though applications are mainly concerned with the case $p = 3$. In particular, the problem of multiple scattering for waves (or particles) in a random medium can be studied with the random processes presented in this work. As originally introduced in [5], compound/mixture processes model allow the description of the output distribution of the direction of propagation of the wave in terms of a mixture of symmetrical distributions on hyperspheres. Similar models are studied in [6], [7] where multiple scattering is described as a Compound Poisson Process (CPP) on the rotation group $SO(3)$. In [6], it is shown that this model allows to describe forward multiple scattering, and its accuracy is high when the number of diffusion events is low. Thus, the CPP model describes the behavior of particles in a scattering medium before the fully developed diffusive regime (known to be thoroughly described by the Brownian motion on $SO(3)$ and originally studied by Perrin [8]).

Convolution on the hypersphere have been studied for *pdf* modelling in engineering application [9]. Random walk models on the sphere have been used to describe wave propagation in engineering literature [10] and multiple scattering is a recurrent issue in many engineering applications such as optics [11], communications [12] and antennas [13], [14]. The occurrence of random scatterers in a wide range of wavelengths makes the problem of multiple scattering relevant to many types of wave propagation. Being able to analyse the distribution of the multiply scattered wave/particle is thus of critical importance. The model studied in this paper aims at describing multiple scattering with a stochastic process and make use of results from harmonic analysis on spheres to predict the behaviour of multiple scattering processes. In addition, stochastic process model allows to infer on the medium the particle/wave has travelled through. In this paper, the harmonic expansions we have derived allows one to numerically evaluate the lower bounds achievable for the estimation of medium parameters.

In this paper, we extend the CPP model to the case where the counting process is no longer a homogeneous Poisson process, but rather a Cox process, *i.e.* a process with intensity being a positive random process itself. The family of processes considered are thus Compound Cox Processes (CCP) on \mathbb{S}^{p-1} . In contrast with [6], [7], the pdf of the random walk and the multiple scattering process is here studied in detail for the general case of isotropic random scattering events and when these events are von Mises Fisher distributed. Several results about multiconvolution, symmetry and unimodality of such pdfs on \mathbb{S}^{p-1} are introduced and used to provide Fourier series expansion of the pdf of a multiply scattered unit vector in \mathbb{R}^p . In the von

N. Le Bihan is with the CNRS, University of Melbourne, Australia. email: nicolas.le-bihan@gipsa-lab.grenoble-inp.fr. His research was supported by the ERA, European Union, through the International Outgoing Fellowship (IOF GeoSToSip 326176) program of the 7th PCRD.

F. Chatelain is with the GIPSA-Lab, Department of Images and Signal, Grenoble, France. email: florent.chatelain@gipsa-lab.grenoble-inp.fr.

J.H. Manton is with the University of Melbourne, Australia. email: jmanton@unimelb.edu.au

Mises Fisher case, we provide an asymptotic distribution of the process which is a mixture of vMF distributions. In addition, we compute the Cramer Rao lower bounds (CRLB) for some parameters of the CCP model on \mathbb{S}^{p-1} in the case where the counting process is either a Poisson process or a Cox process with distribution belonging to an exponential family (when its intensity process is a Gamma process).

The contributions of the paper can be summarized as follows: the isotropic multiple scattering process is expressed using multiple convolution on double cosets. An unimodality theorem is given for such multi-convolution of unimodal and symmetric *pdfs* on \mathbb{S}^{p-1} . Using harmonic analysis results on hyperspheres allows us to obtain the expression of the Fourier coefficients (Legendre polynomial moments) of the *pdf* of a n -step isotropic random walk on \mathbb{S}^{p-1} , which leads to the Fourier expansion of this *pdf*. These results are extended to the multiple scattering process *pdf* when the occurrence of scattering events is a Cox process. In particular, the case when each random step follows a von Mises Fisher (vMF) law is studied in detail: asymptotic approximations for vMF random walk and multiple scattering process on \mathbb{S}^{p-1} are given.

The remainder of this text is outlined as follows. Properties of the isotropic random walk on \mathbb{S}^{p-1} are given in II. Section III presents the Compound Cox Process (CCP) model for the study of multiple scattering on hyperspheres, with emphasis on the use of harmonic analysis on \mathbb{S}^{p-1} to provide characteristic function of the distribution after a time t . Section IV gives an approximation result for multiconvolved vMF *pdfs* and its potential use for the estimation of the CCP parameters.

II. RANDOM WALK ON \mathbb{S}^{p-1}

After reviewing some known facts about hyperspheres and functions taking values on hyperspheres, we introduce new results for the homogeneous random walk on \mathbb{S}^{p-1} .

A. General properties

In \mathbb{R}^p , the hypersphere, denoted \mathbb{S}^{p-1} , is the set of p dimensional vectors with unit length $\mathbb{S}^{p-1} = \{\mathbf{x} \in \mathbb{R}^p; \|\mathbf{x}\| = 1\}$. Hyperspheres (sometimes simply called spheres) are well-known compact manifolds with positive curvature. They are homogeneous spaces of importance in Lie group theory, especially because of their relation with the rotation group $SO(p)$; as they are the following quotients: $\mathbb{S}^{p-1} \cong SO(p)/SO(p-1)$. Hyperspheres are Riemannian symmetric spaces for which the Riemannian distance is simply the "angle" between two elements, $d(\mathbf{x}, \mathbf{y}) = |\arccos(\mathbf{x}^T \mathbf{y})|$ for $\mathbf{x}, \mathbf{y} \in \mathbb{S}^{p-1}$. In the sequel, we will make use of the notation $\theta_{\mathbf{x}, \mathbf{y}}$ for the distance between \mathbf{x} and \mathbf{y} , which is comprised between 0 and π , to avoid any ambiguity. Also, we will use the notation $\mu^T \mathbf{x} = \cos \theta_{\mathbf{x}, \mu}$ as \mathbf{x} and μ are unit vectors in \mathbb{R}^p . Finally, this allows us to define the tangent-normal decomposition of the random unit vector \mathbf{x} :

$$\mathbf{x} = t\mu + \sqrt{1 - t^2}\xi, \quad (1)$$

where $t = \mu^T \mathbf{x}$, and ξ belongs to the intersection of \mathbb{S}^{p-1} with the hyperplane through the origin normal to μ , denoted as $\mu^\perp \cap \mathbb{S}^{p-1}$, which equals \mathbb{S}^{p-2} .

B. Mathematical problem statement

We consider the problem of modelling the distribution of the isotropic multiple scattering process on the sphere \mathbb{S}^{p-1} in \mathbb{R}^p . Each isotropic scattering event acts as a random rotation $R \in SO(p)$ on the direction of propagation $\mathbf{x} \in \mathbb{S}^{p-1}$. The direction after $k \geq 1$ scattering events reads

$$\mathbf{x}_k = R_k \mathbf{x}_{k-1} = R_k \dots R_1 \mathbf{x}_0,$$

where R_l is the rotation matrix associated with the l th scattering event. Let ξ_k be the direction of the normal part of \mathbf{x}_k with respect to \mathbf{x}_{k-1} as defined in (1). The isotropy assumption involves that ξ_k is uniformly distributed on $\mathbf{x}_{k-1}^\perp \cap \mathbb{S}^{p-1}$. Moreover, the scattering events are assumed to be mutually independent and independent of the initial direction \mathbf{x}_0 . This involves the first order Markov property on the chain of random directions $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$ in \mathbb{S}^{p-1} : \mathbf{x}_k given \mathbf{x}_{k-1} is independent of \mathbf{x}_l for $0 \leq l \leq k-2$.

The multiple scattering process considered in this paper is governed by the distribution of $\mathbf{x}_t \equiv \mathbf{x}_{N(t)} \in \mathbb{S}^{p-1}$ where the random number $N(t)$ of scatterers/rotations after a time t is driven by a Poisson or more generally a Cox process.

C. Symmetrical *pdfs* on \mathbb{S}^{p-1}

In this paper, we will consider probability density functions (*pdfs*) f of the direction vector \mathbf{x} . These *pdfs* are elements from $L^1(\mathbb{S}^{p-1}, \mathbb{R})$ with the following additional constraints: positivity and $\int_{\mathbb{S}^{p-1}} f(\mathbf{x}) d\mathbf{x} = 1$. In particular, we will be concerned with *pdfs* that will only depend on the angular variable (the Riemannian distance introduced in II-A). An example of such *pdf* is the von Mises Fisher [15] distribution that will be considered later in the paper. Before moving to this specific case, we consider general symmetrical *pdfs*. In this case, the *pdf* $f(\mathbf{x}; \mu)$ is only a function of the cosine $\mathbf{x}^T \mu$, i.e. for all $\mathbf{x} \in \mathbb{S}^{p-1}$

$$f(\mathbf{x}; \mu) = g(\mu^T \mathbf{x}), \quad (2)$$

where the unit vector parameter $\boldsymbol{\mu} \in \mathbb{S}^{p-1}$ characterizes the rotational axis [15, p. 179]. Without loss of generality, we can assume in the sequel that $\boldsymbol{\mu}$ is oriented such that $E[\boldsymbol{\mu}^T \mathbf{x}] \geq 0$. Consider the tangent-normal decomposition $\mathbf{x} = t\boldsymbol{\mu} + \sqrt{1-t^2}\boldsymbol{\xi}$ defined in (1). The rotational symmetry constraint (2) directly implies that $\boldsymbol{\xi}$ is independent of t and uniformly distributed on the space $\boldsymbol{\mu}^\perp \cap \mathbb{S}^{p-1}$. For symmetry reason, the first moment of area of this space is the null vector. Thus $E[\boldsymbol{\xi}] = \mathbf{0}$, and the mean of \mathbf{x} expresses as

$$E[\mathbf{x}] = \rho\boldsymbol{\mu}, \quad (3)$$

where the scalar $\rho = E[\boldsymbol{\mu}^T \mathbf{x}] = \|E[\mathbf{x}]\| \in [0, 1]$ is called the *mean resultant length* [15, p. 164]. When $\rho > 0$, the *mean direction* is uniquely defined as the rotational axis vector $\boldsymbol{\mu} \in \mathbb{S}^{p-1}$. Note that the mean resultant length ρ is directly linked with the dispersion of the directional distribution. A value of ρ close to 1 indicates a high concentration about the mean direction.

Based now on the higher moments of directional statistics, harmonic analysis on spheres provides us a way to derive a characteristic function for *pdfs* taking values on \mathbb{S}^{p-1} . In the symmetrical case (also known as the *zonal* case), the characteristic function (*i.e.* the Fourier transform of f) takes a simple form [16], [17] as an harmonic basis consists of the Legendre polynomials. Given a *pdf* $f \in L^1(\mathbb{S}^{p-1}, \mathbb{R})$ that is symmetrical about $\boldsymbol{\mu} \in \mathbb{S}^{p-1}$, its characteristic function, denoted \hat{f}_ℓ , for $\ell \geq 0$, is given by:

$$\hat{f}_\ell = E[P_\ell(\cos \theta_{\mathbf{x}, \boldsymbol{\mu}})], \quad (4)$$

$$= E[P_\ell(\boldsymbol{\mu}^T \mathbf{x})], \quad (5)$$

$$= \int_{\mathbb{S}^{p-1}} f(\mathbf{x}; \boldsymbol{\mu}) P_\ell(\boldsymbol{\mu}^T \mathbf{x}) d\mathbf{x} \quad (6)$$

where $P_\ell(\boldsymbol{\mu}^T \mathbf{x})$ are the Legendre polynomials of order ℓ in dimensions p taken at \mathbf{x} with respect to $\boldsymbol{\mu}$, the symmetry axis of f . Note that the Legendre polynomials $P_\ell(t)$ in dimensions p are the same as the ultraspherical or Gegenbauer polynomials $C_\ell^{(p-2)/2}(t)$ [18, pp. 771–802] renormalized such that $P_\ell(1) = 1$. This yields

$$P_\ell(t) = \underbrace{\left[\frac{\Gamma(\ell + p - 2)}{\ell! \Gamma(p - 2)} \right]^{-1}}_{C_\ell^{(p-2)/2}(1)} C_\ell^{(p-2)/2}(t),$$

for all $t \in [-1, 1]$, $\ell \geq 0$. This normalization ensures that $|P_\ell(t)| \leq 1$ for all $|t| \leq 1$, thus $|\hat{f}_\ell| \leq 1$. This family of polynomials forms an orthogonal basis on the Hilbert space of square-integrable functions on \mathbb{S}^{p-1} that are rotationally symmetric about $\boldsymbol{\mu}$:

$$\langle P_\ell, P_m \rangle = \int_{\mathbb{S}^{p-1}} P_\ell(\boldsymbol{\mu}^T \mathbf{x}) P_m(\boldsymbol{\mu}^T \mathbf{x}) d\mathbf{x} = c_{p,\ell}^{-1} \delta_{\ell,m},$$

where $\delta_{\ell,m} = 1$ if $\ell = m$, and 0 otherwise. The normalizing constants read

$$c_{p,\ell} = \frac{1}{\omega_{q-1}} \frac{(2\ell + p - 2)\Gamma(\ell + p - 2)}{\ell! \Gamma(p - 1)}, \quad (7)$$

for all $\ell \geq 0$ with $\omega_{q-1} = 2 \frac{\pi^{p/2}}{\Gamma(p/2)}$ the area of the $(p-1)$ -dimensional sphere \mathbb{S}^{p-1} . Moreover, the Fourier expansion of any p -dimensional rotationally symmetric and continuous *pdf* $f(\cdot; \boldsymbol{\mu})$ can be written:

$$f(\mathbf{x}; \boldsymbol{\mu}) = \sum_{\ell \geq 0} c_{p,\ell} \hat{f}_\ell P_\ell(\boldsymbol{\mu}^T \mathbf{x}), \quad (8)$$

for all $\mathbf{x} \in \mathbb{S}^{p-1}$. Note that the Fourier coefficients \hat{f}_ℓ , also called *Legendre polynomial moments*, are scalar valued and that is a consequence of the symmetry assumption.

D. Convolution of symmetrical pdfs on \mathbb{S}^{p-1}

As already mentioned, we will focus on *pdfs* which are symmetrical and thus only depend on one angular variable. In particular, this will come from the fact that we will consider isotropic random walks on \mathbb{S}^{p-1} . As we will also assume the random steps to be independent, we will end up considering multiple convolution of their associated density. Following [19], we provide here a way to handle the convolution of symmetrical functions in the framework of integrable functions over double cosets. First, recall that given a group G and two subgroups of G , denoted H and K , a double coset in G is an equivalence class defined by the equivalence relation $x \sim y$ iff there exists $h \in H$ and $k \in K$ such that:

$$h x k = y. \quad (9)$$

Given $g \in G$, the double coset $HgK = \{h g k \mid h \in H, k \in K\}$ is therefore the orbit of the group action of $H \times K$ on g , where H acts by left multiplication and K acts by right multiplication. The set of double cosets denoted as $H \backslash G / K$ contains all the orbits

of the group action of $H \times K$ on G . Here, we will consider the case where $G = SO(p)$ and $H = K = SO(p-1)$, i.e. the double cosets in $SO(p-1) \backslash SO(p) / SO(p-1)$. As explained in [19], this set of double cosets can be parametrized using the colatitude measured with respect to the axis left invariant by the $SO(p-1)$ subgroup. The space $L^1(SO(p-1) \backslash SO(p) / SO(p-1), \mathbb{R})$ is the space of functions in $L^1(SO(p), \mathbb{R})$ that are invariant on double cosets KgK for $g \in SO(p)$ and $K \cong SO(p-1)$.

Thus, such functions can be thought as function $g(\mathbf{x}^T \boldsymbol{\mu}) = g(\cos \theta_{\mathbf{x}})$ of the (co)latitude of the $(p-2)$ -dimensional sphere defined by $\cos \theta_{\mathbf{x}} = \mathbf{x}^T \boldsymbol{\mu}$ where $\boldsymbol{\mu} \in \mathbb{S}^{p-1}$ is the axis left invariant by the $SO(p-1)$ rotation subgroup (to be chosen arbitrarily)¹. This underlines that symmetrical *pdfs* about $\boldsymbol{\mu}$ belong to this double coset space according to eq. (2).

The convolution product in $L^1(SO(p-1) \backslash SO(p) / SO(p-1), \mathbb{R})$ is inherited from the convolution product in $L^1(SO(p), \mathbb{R})$ and reads:

$$(f \star_{\boldsymbol{\mu}} g)(\mathbf{x}) = \int_{\mathbb{S}^{p-1}} f(\mathbf{x}^T \mathbf{y}) g(\mathbf{y}^T \boldsymbol{\mu}) d\mathbf{y}, \quad (10)$$

for all $\mathbf{x} \in \mathbb{S}^{p-1}$, where $\boldsymbol{\mu}$ is again the axis left invariant by the rotation subgroup $SO(p-1)$. The notation $\star_{\boldsymbol{\mu}}$ is used to recall that the (co)latitude is measured with respect to the axis $\boldsymbol{\mu}$ left invariant by $SO(p-1)$.

Proposition II.1 (Convolution). *Let $f, g \in L^1(SO(p-1) \backslash SO(p) / SO(p-1), \mathbb{R})$, with $\boldsymbol{\mu}$ standing for the axis left invariant by the rotation subgroup $SO(p-1)$. The following properties hold for the convolution in the double coset space.*

i) *Stability.* $f \star_{\boldsymbol{\mu}} g \in L^1(SO(p-1) \backslash SO(p) / SO(p-1), \mathbb{R})$, that is $f \star_{\boldsymbol{\mu}} g$ is a function of the only cosine $\boldsymbol{\mu}^T \mathbf{x}$:

$$(f \star_{\boldsymbol{\mu}} g)(\mathbf{x}) \equiv (f \star_{\boldsymbol{\mu}} g)(\boldsymbol{\mu}^T \mathbf{x}), \quad (11)$$

ii) *Commutativity.*

$$(f \star_{\boldsymbol{\mu}} g)(\mathbf{x}) = (g \star_{\boldsymbol{\mu}} f)(\mathbf{x}), \quad (12)$$

iii) *Fourier product.* If \widehat{f}_{ℓ} and \widehat{g}_{ℓ} are the respective ℓ th-order Fourier coefficients of f and g , the Fourier coefficients of their convolution product is given by:

$$\widehat{(f \star_{\boldsymbol{\mu}} g)}_{\ell} = \widehat{f}_{\ell} \widehat{g}_{\ell}, \quad (13)$$

for all $\ell \geq 0$.

Proof: Some proofs for the two first properties i) and ii) are given in [19, p. 237–239] for $p = 3$; they naturally extend to the general case $p \geq 2$. Property iii) is a direct consequence of the Funk-Hecke theorem [17, Theorem 7.8, p. 188]. ■

These properties illustrate that the convolution product and its Fourier expansion behave nicely in this space. In particular, property iii) of Prop. II.1 allows us to obtain the convolution theorem. Finally, some well-known results on symmetry and unimodality in the real line can be extended to the convolution on the hypersphere.

Theorem II.2 (Unimodality of the convolution product). *Let $f, g \in L^1(SO(p-1) \backslash SO(p) / SO(p-1), \mathbb{R})$ be the *pdfs* of two absolutely continuous unimodal and rotationally symmetric distributions on \mathbb{S}^{p-1} with the same mode which necessarily equals their mean direction $\boldsymbol{\mu} \in \mathbb{S}^{p-1}$. Then the convolved distribution $f \star_{\boldsymbol{\mu}} g$, which is rotationally symmetric about $\boldsymbol{\mu}$ or equivalently belongs to $L^1(SO(p-1) \backslash SO(p) / SO(p-1), \mathbb{R})$ according to Prop. II.1, is also unimodal with mode $\boldsymbol{\mu}$.*

Proof: See Appendix A. ■

E. Random walk on \mathbb{S}^{p-1} : directional distribution

As explained in II-B, the chain of random vectors $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$ in \mathbb{S}^{p-1} obeys the Markov property since given all the past directions $\mathbf{x}_0, \dots, \mathbf{x}_{k-1}$, the current direction \mathbf{x}_k depends only on the previous one \mathbf{x}_{k-1} . As a consequence, each random step $\mathbf{x}_{k-1} \rightarrow \mathbf{x}_k$, for all $k \geq 1$, are independent. This defines a discrete time random walk on the hypersphere \mathbb{S}^{p-1} . It is of note that these steps are not necessarily identically distributed. However, an important case appears when they are isotropic, so that all the step directions are equiprobable. Thus, the distribution of the k th-step direction \mathbf{x}_k is rotationally symmetric about the previous one \mathbf{x}_{k-1} , for all $k \geq 1$. According to (2), its conditional *pdf* expresses as

$$f(\mathbf{x}_k | \mathbf{x}_{k-1}) = g_{k,k-1}(\mathbf{x}_{k-1}^T \mathbf{x}_k), \quad (14)$$

for all $k \geq 1$. In the remainder, we assume that the initial direction \mathbf{x}_0 is fixed to a deterministic direction $\boldsymbol{\mu}$ and that the walk is isotropic, which means that the distributions that govern each step express as (14). The following proposition states

¹A function $h \in L^1(SO(p-1) \backslash SO(p) / SO(p-1), \mathbb{R})$ is defined on the rotation group $SO(p)$. However, this function depends only on the cosine between the axis left invariant by the $SO(p-1)$ subgroup and its rotated image. For any $R \in SO(p)$, such a function can be expressed as $h(R) = f(\mathbf{x}) = g(\boldsymbol{\mu}^T \mathbf{x})$ with $\mathbf{x} = R\boldsymbol{\mu}$, where the functions f and g are defined on \mathbb{S}^{p-1} and $[-1, 1]$ respectively. This shows that a function in $L^1(SO(p-1) \backslash SO(p) / SO(p-1), \mathbb{R})$ can be identified as a function on \mathbb{S}^{p-1} or a function of the cosine $\boldsymbol{\mu}^T \mathbf{x}$ defined on $[-1, 1]$. By abuse of notation, we will confuse these functions in the remainder.

the link between the directional *pdf* after n steps of the isotropic random walk on \mathbb{S}^{p-1} with the convolution on the double coset $L^1(SO(p-1) \backslash SO(p)/SO(p-1), \mathbb{R})$ introduced in section II-D.

Theorem II.3. *Given an isotropic n -step random walk on \mathbb{S}^{p-1} , the pdf of $\mathbf{x}_n \in \mathbb{S}^{p-1}$ is the n -fold convolution in $L^1(SO(p-1) \backslash SO(p)/SO(p-1), \mathbb{R})$, where $SO(p-1)$ is the rotation subgroup such that $\boldsymbol{\mu}$ is left invariant. It reads*

$$f(\mathbf{x}_n; \boldsymbol{\mu}) = (g_{n,n-1} \star_{\boldsymbol{\mu}} \cdots \star_{\boldsymbol{\mu}} g_{1,0})(\mathbf{x}_n) \quad (15)$$

where $f(\mathbf{x}_k | \mathbf{x}_{k-1}) = g_{k,k-1}(\mathbf{x}_k^T \mathbf{x}_{k-1})$ can be identified as the conditional pdf of \mathbf{x}_k given \mathbf{x}_{k-1} .

Proof: The proof is conducted by induction. Since $\mathbf{x}_0 = \boldsymbol{\mu}$ is a deterministic vector, the pdf of \mathbf{x}_1 is $f(\mathbf{x}_1; \boldsymbol{\mu}) = g_{1,0}(\boldsymbol{\mu}^T \mathbf{x}_1)$ and belongs to $L^1(SO(p-1) \backslash SO(p)/SO(p-1), \mathbb{R})$. Thus the base case holds for $n = 1$. Assume now that the pdf of \mathbf{x}_{k-1} , for $k > 1$, is symmetrical with respect to $\boldsymbol{\mu}$, i.e. $f(\mathbf{x}_{k-1}; \boldsymbol{\mu}) = g_{k-1}(\boldsymbol{\mu}^T \mathbf{x}_{k-1}) \in L^1(SO(p-1) \backslash SO(p)/SO(p-1), \mathbb{R})$, and is given by the following $(k-1)$ -fold convolution: $f_{k-1}(\mathbf{x}_{k-1}; \boldsymbol{\mu}) = (g_{k-1,k-2} \star_{\boldsymbol{\mu}} \cdots \star_{\boldsymbol{\mu}} g_{1,0})(\mathbf{x}_{k-1})$.

Due to the isotropic assumption, the conditional pdf $f(\mathbf{x}_k | \mathbf{x}_{k-1}) = g_{k,k-1}(\mathbf{x}_k^T \mathbf{x}_{k-1})$ also belongs to $L^1(SO(p-1) \backslash SO(p)/SO(p-1), \mathbb{R})$. Moreover, this conditional pdf allows us to express the density of \mathbf{x}_k as

$$\begin{aligned} f(\mathbf{x}_k; \boldsymbol{\mu}) &= \int_{\mathbb{S}^{p-1}} f(\mathbf{x}_k | \mathbf{x}_{k-1}) f(\mathbf{x}_{k-1}; \boldsymbol{\mu}) d\mathbf{x}_{k-1}, \\ &= \int_{\mathbb{S}^{p-1}} g_{k,k-1}(\mathbf{x}_k^T \mathbf{x}_{k-1}) g_{k-1}(\boldsymbol{\mu}^T \mathbf{x}_{k-1}) d\mathbf{x}_{k-1}. \end{aligned}$$

According to (10), we recognize the following convolution on the double coset: $f(\mathbf{x}_k; \boldsymbol{\mu}) = (g_{k,k-1} \star_{\boldsymbol{\mu}} g_{k-1})(\mathbf{x}_k)$. Thus $f(\mathbf{x}_k; \boldsymbol{\mu})$ is also symmetrical about $\boldsymbol{\mu}$ according to property i) of Prop. II.1. As g_{k-1} is assumed to be a $k-1$ -fold convolution, it comes finally by associativity that $f(\mathbf{x}_k; \boldsymbol{\mu}) = (g_{k,k-1} \star_{\boldsymbol{\mu}} \cdots \star_{\boldsymbol{\mu}} g_{1,0})(\mathbf{x}_k)$, and the inductive step holds. ■

A direct consequence of Theorem II.3 is that the distribution of n -step random walk is rotationally symmetrical about the initial direction $\boldsymbol{\mu}$ since $f(\mathbf{x}_n; \boldsymbol{\mu}) \in L^1(SO(p-1) \backslash SO(p)/SO(p-1), \mathbb{R})$. Furthermore, it allows us to express the characteristic function of the random walk based on the Legendre polynomial moments of each step.

Corollary II.4 (Mean and Fourier Coefficient of the isotropic random walk). *For all $n \geq 1$, the mean of the n -step direction $\mathbf{x}_n \in \mathbb{S}^{p-1}$ expresses as*

$$E[\mathbf{x}_n] = \left(\prod_{k=1}^n \rho_{k,k-1} \right) \boldsymbol{\mu}, \quad (16)$$

where $\rho_{k,k-1} = E[\mathbf{x}_{k-1}^T \mathbf{x}_k | \mathbf{x}_{k-1}]$ is the mean resultant length for the conditional distribution $f(\mathbf{x}_k | \mathbf{x}_{k-1}) = g_{k,k-1}(\mathbf{x}_{k-1}^T \mathbf{x}_k)$ that governs the k th step $\mathbf{x}_{k-1} \rightarrow \mathbf{x}_k$.

More generally, for all $n \geq 1$, $\ell \geq 0$, the ℓ th order Fourier coefficient of the n -step distribution is

$$\widehat{f}_\ell^{\otimes n} \equiv E[P_\ell(\boldsymbol{\mu}^T \mathbf{x}_n)] = \prod_{k=1}^n \widehat{g_{k,k-1}}_\ell, \quad (17)$$

where $\widehat{g_{k,k-1}}_\ell = E[P_\ell(\mathbf{x}_{k-1}^T \mathbf{x}_k) | \mathbf{x}_{k-1}]$ denotes the Fourier coefficient for the conditional distribution $f(\mathbf{x}_k | \mathbf{x}_{k-1}) = g_{k,k-1}(\mathbf{x}_{k-1}^T \mathbf{x}_k)$.

Proof: Eq. (17) is derived from the multiconvolution formula (15) and the iterative use of the convolution theorem (13) given in Prop. II.1. Eq. (16) is derived from the mean of a rotationally symmetric distribution (3) and from (17) when $\ell = 1$ since $P_1(t) = t$ for all $p \geq 2$. ■

Eq. (16) shows that the mean direction of the n -step is the initial direction $\boldsymbol{\mu}$, while its mean resultant length reduces to the product of the mean resultant lengths associated with each step:

$$\rho_n \equiv E[\boldsymbol{\mu}^T \mathbf{x}_n] = \prod_{k=1}^n \rho_{k,k-1}. \quad (18)$$

This formula emphasizes that the directional dispersion increases with the number n of steps, since $0 \leq \rho_{k,k-1} \leq 1$ for all $k \geq 1$.

Based on the Fourier coefficient formulas, it becomes possible to obtain a Fourier expansion of the rotationally symmetric random walk pdf.

Corollary II.5 (Fourier expansion of the isotropic random walk pdf). *For all $n \geq 1$, the Fourier expansion pdf of the n -step direction $\mathbf{x}_n \in \mathbb{S}^{p-1}$ reads*

$$f(\mathbf{x}_n; \boldsymbol{\mu}) = \sum_{\ell \geq 0} c_{p,\ell} \widehat{f}_\ell^{\otimes n} P_\ell(\boldsymbol{\mu}^T \mathbf{x}), \quad (19)$$

where the Fourier coefficients $\widehat{f}_\ell^{\otimes n}$ are given in (17), the normalizing constants $c_{p,\ell}$ being defined in (7).

Proof: The distribution of \mathbf{x}_n is symmetric according to Theorem II.3, and the Fourier expansion formula (8) can be applied. ■

It is important to note that when all the steps of the random walk are identically distributed, i.e. $f(\mathbf{x}_k|\mathbf{x}_{k-1}) = g(\mathbf{x}_{k-1}^T \mathbf{x}_k)$ for all $k \geq 1$, the Fourier coefficients reduces to

$$\hat{f}_\ell^{\otimes n} = (\hat{g}_\ell)^n, \quad (20)$$

where $\hat{g}_\ell = E[P_\ell(\boldsymbol{\mu}^T \mathbf{x}_1)]$ is the Fourier coefficient of the distribution that governs a random walk step.

Finally, it is possible to derive sufficient conditions to ensure the unimodality of the isotropic random walk

Corollary II.6 (Unimodality of the isotropic random walk). *Assume that the conditional directional distributions $f(\mathbf{x}_k|\mathbf{x}_{k-1}) = g_{k,k-1}(\mathbf{x}_{k-1}^T \mathbf{x}_k)$, which are rotationally symmetric, are also absolutely continuous (or equivalently that $g_{k,k-1}$ is a continuous function on $[-1, 1]$) and unimodal with a mode that necessarily equals their mean direction. Then, for all $n \geq 1$, the pdf of the n -step direction $\mathbf{x}_n \in \mathbb{S}^{p-1}$ is also unimodal with a mode that equals the original direction $\boldsymbol{\mu} \equiv \mathbf{x}_0$.*

Proof: This result is derived from the multiconvolution formula (15) and the iterative use of Theorem II.2. ■

III. MULTIPLE SCATTERING MODEL

The model presented in this section is motivated by the description of multiple scattering which occurs in a wide range of applications in Physics and Engineering, including optical, microwave, acoustics or elastic waves [20]. We consider the description of the distribution of the output direction of propagation of a particle/wave that propagated through a random medium. This random medium is made of an homogeneous medium/matrix containing some inclusions of size of the same order as the particle size (or wavelength of the wave). Inclusions have different physical properties inducing that a scattering event happens each time the particle/wave encounters an inclusion (also named scatterer). The number and locations of the scatterers is random and between two scattering events, the wave/particle propagates ballistically. It is a classical approach to consider that the time between two scattering events follows an exponential law. Such an assumption leads to Compound Poisson Process models as described in [5], [6], [7], [21]. Here, we consider the more general case where the intensity of the counting process is a random process itself. We will thus make use of Compound Cox Processes taking values on hyperspheres to model multiple scattering.

A. Compound Cox process on \mathbb{S}^{p-1}

Consider an initial vector $\mathbf{x}_0 \equiv \boldsymbol{\mu} \in \mathbb{S}^{p-1}$. After a time t (the time spent propagating in the random medium), assume the resulting vector $\mathbf{x}_t \in \mathbb{S}^{p-1}$ is a mixture made of contributions of rotated versions of $\boldsymbol{\mu}$ an arbitrary number of times n . The weight of each contribution is simply the probability that the wave/particle encountered n scatterers during the period of time t , i.e. $\mathbb{P}[N(t) = n]$. $N(t)$ is called the *counting process*. In the classical compound Poisson process [7], $N(t)$ is an homogeneous Poisson process and each individual weight is equal to $e^{-\lambda t} (\lambda t)^n / n!$ where the constant λ is the Poisson intensity parameter. This weight is obtained when the time between two rotations of the vector is chosen to have an exponential distribution with parameter λ . The equivalent Poisson parameter $\lambda_t = \lambda t$ of $N(t)$ consists of the mean number of rotation events in the elapsed time t . In Physics, it is related to the *mean free path* ℓ like $\ell = c/\lambda$ where c is the celerity in the medium. Thus, ℓ is the mean distance between two consecutive rotation events (see [6]).

Now, if the counting process $N(t)$ is no more a homogeneous Poisson process, an alternative is to consider that the intensity measure of $N(t)$ is a random process $\Lambda(t)$. $N(t)$ is then called a mixed Poisson process, or a Cox process [22].

In this case, the distribution of $N(t)$ is a mixed Poisson distribution which reads [23]:

$$\mathbb{P}[N(t) = n] = \mathcal{P}_n[f_{\Lambda(t)}] \quad (21)$$

$$= \int_0^{+\infty} \frac{e^{-\lambda_t} \lambda_t^n}{n!} f_{\Lambda(t)}(\lambda_t) d\lambda_t \quad (22)$$

where $\mathcal{P}_n[f_{\Lambda(t)}]$ is called the *Poisson transform* of the mixing pdf $f_{\Lambda(t)}$ [23].

In the isotropic case, the steps associated with the scattering events are governed by rotationally symmetric distribution as explained in section II-E. For the sake of simplicity, all these random steps are assumed to be identically distributed, and we denote as $\hat{f}_\ell \equiv \hat{g}_\ell$ the ℓ th order Fourier coefficient of the random step distribution, for all $\ell \geq 0$. This yields that the ℓ th order Fourier coefficient of the n -step random walk direction reduces to $\hat{f}_\ell^{\otimes n} = (\hat{f}_\ell)^n$ according to (20).

Conditioning by the number of scattering events, one gets the expression of the density of \mathbf{x}_t :

$$\begin{aligned} f(\mathbf{x}_t; \boldsymbol{\mu}) &= \mathcal{P}_0[f_{\Lambda_t}] \delta_{\boldsymbol{\mu}}(\mathbf{x}_t) \\ &+ \sum_{n \geq 1} \mathcal{P}_n[f_{\Lambda_t}] f^{\otimes n}(\mathbf{x}_t; \boldsymbol{\mu}), \end{aligned} \quad (23)$$

where $\delta_{\mu}(\mathbf{x}_t)$ denotes a mass located in the original direction $\mu \equiv \mathbf{x}_0 \in \mathbb{S}^{p-1}$, and $f^{\otimes n}(\mathbf{x}_t; \mu)$ denotes the n -step random walk pdf with original direction μ .

The pdf $f(\mathbf{x}_t; \mu)$ thus consists of a mixture of n -fold convolutions of identical distributions according to Thm. II.3, for all $n > 0$, plus a mass in μ which corresponds to direct paths.

Equation (23) covers several cases. If $N(t)$ is a homogeneous Poisson process with intensity parameter λ , then $\mathcal{P}_n[f_{\Lambda_t}(w_t)] = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$ and this case was considered in [21], [6], [7]. If $N(t)$ is a Cox process, an interesting case appears when its distribution belongs to an exponential family. As explained in [24], this happens when $\Lambda(t)$ is a stationary i.i.d. Gamma process, i.e. $\Lambda_t \sim \mathcal{G}(t; \xi, \theta) \sim \mathcal{G}(\xi t, \theta)$ where ξ is the shape parameter and θ the scale parameter. The distribution of this process is thus:

$$f_{\Lambda_t}(\lambda_t = x) = \frac{\theta^{\xi t}}{\Gamma(\xi t)} x^{\xi t - 1} e^{-\theta x}, \quad (24)$$

for all $x > 0$. It comes by direct calculation that in this case the Poisson transform is

$$\mathcal{P}_n[f_{\Lambda_t}] = \frac{\Gamma(n + \xi t)}{n! \Gamma(\xi t)} \frac{\theta^{\xi t}}{(\theta + 1)^{n + \xi t}} \quad (25)$$

which shows that when Λ_t is a Gamma process with scale parameter ξ and shape parameter θ , $N(t)$ is a negative binomial process. In fact, in such case, the weight coefficients in (23) follow a negative binomial law $\mathcal{NB}(r_t, q)$ with stopping-time parameter $r_t = \xi t$ and success probability $q = (\theta + 1)^{-1}$.

B. Characteristic function of the multiple scattering process

The distribution of the multiple scattering process (23) consists of the mixture of a mass in μ for the direct paths and a continuous distribution that consists of the n -fold random step convolutions for all $n \geq 1$. This continuous distribution is the conditional distribution of \mathbf{x}_t given there is at least one diffusion, i.e. $N(t) > 0$. Its pdf denoted as $f^{\otimes > 0}$ reads $f^{\otimes > 0}(\mathbf{x}_t; \mu) = c_0 h^{\otimes > 0}(\mathbf{x}_t; \mu)$ where $c_0 = (1 - \mathcal{P}_0[f_{\Lambda_t}])^{-1}$ is the normalizing constant of the truncated distribution corresponding to the event $N(t) > 0$ and $h^{\otimes > 0}$ is the following unnormalized density

$$h^{\otimes > 0}(\mathbf{x}_t; \mu) = \sum_{n \geq 1} \mathcal{P}_n[f_{\Lambda_t}] f^{\otimes n}(\mathbf{x}_t; \mu). \quad (26)$$

This pdf is rotationally symmetric about μ as a mixture of symmetric distributions. This shows that $h^{\otimes > 0}(\mathbf{x}_t; \mu)$ admits a Fourier expansion (8). Note also that when the distribution that governs the random steps is unimodal, the pdf of \mathbf{x}_t is unimodal with mode μ as a mixture of unimodal distributions according to Corollary II.6.

Considering the Laplace transform of the mixing process Λ_t

$$\mathcal{L}_{\Lambda_t}[z] = E[e^{-z\Lambda_t}] = \int_0^{+\infty} e^{-z\lambda_t} f_{\Lambda_t}(\lambda_t) d\lambda_t,$$

we obtain the following expression of the Fourier coefficients.

Lemma 1. *The Fourier coefficients of the continuous unnormalized density $h^{\otimes > 0}$ express as*

$$\widehat{h^{\otimes > 0}}_{\ell} = \mathcal{L}_{\Lambda_t} [1 - \widehat{f}_{\ell}] - \mathcal{L}_{\Lambda_t} [1], \quad (27)$$

for all $\ell \geq 0$, where \widehat{f}_{ℓ} is the Fourier coefficient of the isotropic and identically distributed random steps and where $\mathcal{L}_{\Lambda_t} [1] = \mathcal{P}_0[f_{\Lambda_t}] = \Pr(N_t = 0)$.

Proof: Based on the orthogonality property of the Legendre polynomials, it comes from (26) and the Fourier expansion of each n -fold convolved pdf $f^{\otimes n}$ (19), for $n \geq 1$, that

$$\widehat{h^{\otimes > 0}}_{\ell} = \sum_{n \geq 1} \mathcal{P}_n[f_{\Lambda_t}] \widehat{f}_{\ell}^n, \quad (28)$$

where $|\widehat{f}_{\ell}| \leq 1$ by construction. According to Fubini theorem, one can interchange the summation symbol with the Poisson transform integral. This yields

$$\begin{aligned} \widehat{h^{\otimes > 0}}_{\ell} &= \int_0^{+\infty} e^{-\lambda_t} f_{\Lambda_t}(\lambda_t) \sum_{n \geq 1} \frac{(\lambda_t \widehat{f}_{\ell})^n}{n!} d\lambda_t, \\ &= \int_0^{+\infty} e^{-\lambda_t} f_{\Lambda_t}(\lambda_t) [e^{\lambda_t \widehat{f}_{\ell}} - 1] d\lambda_t, \\ &= \mathcal{L}_{\Lambda_t} [1 - \widehat{f}_{\ell}] - \mathcal{L}_{\Lambda_t} [1], \end{aligned}$$

for all $\ell \geq 0$. Note finally that the Laplace transform $\mathcal{L}_{\Lambda_t}[z]$ is well-defined for all $z \geq 0$ since Λ_t is a positive random variable. Thus the Fubini theorem holds as $1 - |\widehat{f}_\ell| \geq 0$, and the Fourier coefficients are well defined. ■

Consider now the probability generating function of the Cox process N_t :

$$G_{N_t}[z] = E[z^{N_t}] = \sum_{n \geq 0} \mathcal{P}_n[f_{\Lambda_t}] z^n.$$

A classical result about mixed Poisson distribution, see for instance [24], [25], is that G_{N_t} can be easily derived from the Laplace transform of the mixing distribution as

$$G_{N_t}[z] = \mathcal{L}_{\Lambda_t}[1 - z].$$

The Fourier coefficients given in Lemma 1 can thus be expressed in an equivalent way as

$$\widehat{h^{\otimes > 0}}_\ell = G_{N_t}[\widehat{f}_\ell] - G_{N_t}[0], \quad (29)$$

where $G_{N_t}[0] = \mathcal{L}_{\Lambda_t}[1] = \mathcal{P}_0[f_{\Lambda_t}] = \Pr(N_t = 0)$.

Finally, based on the Fourier expansion of the continuous function $h^{\otimes > 0}$, we obtain the following results for the distribution of \mathbf{x}_t .

Proposition III.1. *The pdf of the direction \mathbf{x}_t in the multiple scattering process can be expanded as*

$$f(\mathbf{x}_t; \boldsymbol{\mu}) = \mathcal{P}_0[f_{\Lambda_t}] \delta_{\boldsymbol{\mu}}(\mathbf{x}_t) + \sum_{\ell \geq 0} c_{p,\ell} \widehat{h^{\otimes > 0}}_\ell P_\ell(\boldsymbol{\mu}^T \mathbf{x}_t), \quad (30)$$

where the coefficients $\widehat{h^{\otimes > 0}}_\ell$ are given in (27), or equivalently in (29).

Moreover, the Legendre polynomial moments read

$$E[P_\ell(\boldsymbol{\mu}^T \mathbf{x}_t)] = G_{N_t}[\widehat{f}_\ell] = \mathcal{L}_{\Lambda_t}[1 - \widehat{f}_\ell], \quad (31)$$

for all $\ell \geq 0$, where \widehat{f}_ℓ is the Fourier coefficient of the isotropic and identically distributed random steps.

Proof: It remains to show the Legendre polynomial moment formula (31). The distribution of \mathbf{x}_t is a linear mixture of a mass in $\boldsymbol{\mu}$ with probability $\mathcal{P}_0[f_{\Lambda_t}] = G_{N_t}[0]$ and a continuous distribution. As $P_\ell(\boldsymbol{\mu}^T \boldsymbol{\mu}) = P_\ell(1) = 1$, the contribution of the mass in $\boldsymbol{\mu}$ to the ℓ th order Legendre moment reduces to its probability. By linearity, it comes that $E[P_\ell(\boldsymbol{\mu}^T \mathbf{x}_t)] = G_{N_t}[0] + \widehat{h^{\otimes > 0}}_\ell = G_{N_t}[\widehat{f}_\ell]$ according to (29). ■

It is important to note that the distribution of \mathbf{x}_t is not continuous due to the mass in $\boldsymbol{\mu}$. As written in (30), only the continuous part admit a Fourier expansion in the pointwise convergence sense. However, all the Legendre polynomial moments of the multiple scattering direction \mathbf{x}_t are well-defined and have tractable expressions given in (31).

One consequence is that the mean resultant length of the multiple scattering process \mathbf{x}_t express as $\rho_t = E[\boldsymbol{\mu}^T \mathbf{x}_t] = G_{N_t}[\widehat{f}_1]$. And due to the rotational symmetry, the mean of \mathbf{x}_t is

$$E[\mathbf{x}_t] = \rho_t \boldsymbol{\mu} = G_{N_t}[\widehat{f}_1] \boldsymbol{\mu}.$$

Proposition III.1 allows us to express the Fourier coefficients and the Legendre polynomial moments in two particular cases that we are especially interested in. First in the case of the homogeneous Poisson process where $\Lambda_t = \lambda_t$ is a deterministic constant, it comes that

$$\begin{aligned} \widehat{h^{\otimes > 0}}_\ell &= e^{-\lambda_t} \left[e^{\lambda_t \widehat{f}_\ell} - 1 \right], \\ E[P_\ell(\boldsymbol{\mu}^T \mathbf{x}_t)] &= e^{-\lambda_t [1 - \widehat{f}_\ell]}, \end{aligned} \quad (32)$$

for all $\ell \geq 0$. Second, when $\Lambda_t \sim \mathcal{G}(\xi t, \theta)$ is a Gamma process, i.e. that $N_t \sim \mathcal{NB}(r_t, q)$ is a Negative Binomial process with $r_t = \xi t$ and $q = (\theta + 1)^{-1}$, straightforward computations lead to

$$\begin{aligned} \widehat{h^{\otimes > 0}}_\ell &= (1 - q)^{r_t} \left[\left(1 - q \widehat{f}_\ell \right)^{-r_t} - 1 \right], \\ E[P_\ell(\boldsymbol{\mu}^T \mathbf{x}_t)] &= \left(\frac{1 - q}{1 - q \widehat{f}_\ell} \right)^{r_t}, \end{aligned} \quad (33)$$

for all $\ell \geq 0$.

IV. VON MISES-FISHER RANDOM WALK AND MULTIPLE SCATTERING PROCESS ON \mathbb{S}^{p-1}

A. von Mises-Fisher distribution on \mathbb{S}^{p-1}

The von Mises-Fisher distribution [15, p. 167] is probably the most important distribution in the statistics of hyperspherical data and plays a role on \mathbb{S}^{p-1} analogue to the role of the normal distribution on the real line. This distribution, denoted as $M_p(\boldsymbol{\mu}, \kappa)$, is defined by the following *pdf* for all $\mathbf{x} \in \mathbb{S}^{p-1}$

$$f(\mathbf{x}; \boldsymbol{\mu}, \kappa) = \frac{\kappa^{p/2-1}}{(2\pi)^{p/2} I_{p/2-1}(\kappa)} e^{\kappa \mathbf{x}^T \boldsymbol{\mu}}, \quad (34)$$

where $I_\nu(\cdot)$ is the modified Bessel function [18, p. 374], $\boldsymbol{\mu} \in \mathbb{S}^{p-1}$ corresponds to the mean direction and $\kappa \geq 0$ is the *concentration parameter*: the larger the value of κ , the more concentrated is the distribution about the mean direction $\boldsymbol{\mu}$. Conversely, when $\kappa = 0$ the distribution reduces to the uniform distribution on \mathbb{S}^{p-1} . Its mean resultant length takes the form:

$$\rho \equiv A_p(\kappa) = \frac{I_{p/2}(\kappa)}{I_{p/2-1}(\kappa)}, \quad (35)$$

which reduces to

$$\rho \equiv A_3(\kappa) = \coth \kappa - \frac{1}{\kappa},$$

when $p = 3$. Based on the *pdf* expression (34), it is straightforward to see that this distribution is rotationally symmetric and unimodal with mode $\boldsymbol{\mu}$ when $\kappa > 0$. Finally, as explained in [16], the characteristic function of the von Mises Fisher distribution $f(\mathbf{x}; \boldsymbol{\mu}, \kappa)$ takes the form:

$$\hat{f}_\ell(\kappa) = E[P_\ell(\mathbf{x}^T \boldsymbol{\mu})] = \frac{I_{\ell+\nu}(\kappa)}{I_\nu(\kappa)}, \quad (36)$$

where $\nu = p/2 - 1$, for $\kappa > 0$ and $\ell \geq 0$.

B. von Mises-Fisher random walk

One problem of characterizing more deeply and inferring efficiently the distribution of the n th-step direction \mathbf{x}_n is that, except the Fourier series expansion (19), there is no simple closed form expression of the density of the multiply convolved distribution (15).

However, when all the isotropic random walk steps are governed by unimodal distributions, Corollary II.6 says that the n -step random walk direction \mathbf{x}_n is also governed by an unimodal rotationally symmetric distributions with mode the original direction $\boldsymbol{\mu} \equiv \mathbf{x}_0$. This suggests that the distribution of \mathbf{x}_n can be well fitted by a standard unimodal rotationally symmetric distributions with mode $\boldsymbol{\mu}$. Due to the properties of the vMF distributions presented in section IV-A, this family seems to be a good candidate to fit the distribution of \mathbf{x}_n , for $n \geq 1$. It leads to model the distribution of \mathbf{x}_n by a $M_p(\boldsymbol{\mu}, \tilde{\kappa}_n)$ distribution, where $\tilde{\kappa}_n$ is an equivalent concentration parameter for the n -step direction.

Moreover, we consider now that all the random walk steps (14) are identically distributed according to a vMF distribution with concentration parameter κ , that is

$$\mathbf{x}_k | \mathbf{x}_{k-1} \sim M_p(\mathbf{x}_{k-1}, \kappa),$$

and the resulting random walk is called the *vMF random walk* with concentration parameter κ . This random walk is isotropic and unimodal with mode $\boldsymbol{\mu} \equiv \mathbf{x}_0$ according to Corollary II.6. It is possible to obtain, in the high concentration case, a simple vMF asymptotic distribution for the n -step direction.

Theorem IV.1. *Consider the vMF random walk with concentration parameter κ . Then, in the large κ and small $n \geq 1$ case, i.e. when $n/\kappa \rightarrow 0$, \mathbf{x}_n is asymptotically distributed as $M_p(\boldsymbol{\mu}, \tilde{\kappa}_n)$ where*

$$\tilde{\kappa}_n = \frac{\kappa - 1/2}{n} + 1/2 \quad (37)$$

is the equivalent concentration parameter. The asymptotic distribution yields a third-order approximation of the Fourier coefficients

$$\hat{f}_\ell^{\otimes n} = \tilde{f}_\ell^n + O\left(\left(\frac{n}{\kappa}\right)^3\right), \quad \text{as } \frac{n}{\kappa} \rightarrow 0,$$

for any $\ell \geq 0$, where $\hat{f}_\ell^{\otimes n}$ is the ℓ th-order Fourier coefficient of the \mathbf{x}_n distribution, while \tilde{f}_ℓ^n denotes the ℓ th-order Fourier coefficient of the asymptotic distribution $M_p(\boldsymbol{\mu}, \tilde{\kappa}_n)$.

Proof: Both $M_p(\boldsymbol{\mu}, \tilde{\kappa}_n)$ and the distribution of \mathbf{x}_n are rotationally symmetric about $\boldsymbol{\mu}$. Thus it is sufficient to show the asymptotic equivalence of their tangent part about $\boldsymbol{\mu} \in \mathbb{S}^{p-1}$. The tangent part is bounded in $[-1, 1]$, which ensures that its

moments are well-defined and also belong to $[-1, 1]$. Thus its distribution is uniquely defined by its moments, see for instance [26, Theorem 30.1, p. 388]. As the family of Legendre polynomials form a polynomial basis, we can conclude by the method of moments that the distributions are asymptotically equivalent if their Fourier coefficients are asymptotically equivalent.

For $n \geq 1$, we obtain from (20) and (36), and from (36) and (37) respectively, that

$$\widehat{f}_\ell^{\otimes n} = \left(\frac{I_{\ell+\nu}(\kappa)}{I_\nu(\kappa)} \right)^n, \quad \widetilde{f}_\ell^n = \frac{I_{\ell+\nu}(\tilde{\kappa}_n)}{I_\nu(\tilde{\kappa}_n)}.$$

Using now the following asymptotic expansion of the modified Bessel function for large κ [18, p. 377]:

$$I_\nu(\kappa) = \frac{e^\kappa}{\sqrt{2\pi\kappa}} \left[1 - \frac{4\nu^2 - 1}{8\kappa} + \frac{(4\nu^2 - 1)(4\nu^2 - 9)}{2!(8\kappa)^2} + O\left(\frac{1}{\kappa^3}\right) \right],$$

yields that, for any $\ell \geq 0$,

$$\widehat{f}_\ell^{\otimes n} = 1 - \frac{\ell n(\ell + 2\nu)}{2\kappa} + \frac{\ell n(\ell + 2\nu)(n\ell^2 + 2n\nu\ell - 2)}{8\kappa^2} + O\left(\frac{n^3}{\kappa^3}\right),$$

and that

$$\widetilde{f}_\ell^n = 1 - \frac{\ell n(\ell + 2\nu)}{2\kappa} + \frac{\ell n(\ell + 2\nu)(n\ell^2 + 2n\nu\ell - 2)}{8\kappa^2} + O\left(\frac{n^3}{\kappa^3}\right).$$

This shows that $\widehat{f}_\ell^{\otimes n} = \widetilde{f}_\ell^n + O\left(\left(\frac{n}{\kappa}\right)^3\right)$ as $\frac{n}{\kappa}$ tends to zero, for any $\ell \geq 0$, which concludes the proof. \blacksquare

Note that the expression (37) of the equivalent concentration parameter has been derived in [21] by matching the mean resultant length in the asymptotic case. Theorem IV.1 shows that a similar result extends to all the Legendre polynomial moments, and thus to the distribution.

To appreciate the accuracy of the vMF approximation given by Theorem IV.1, Fig. 1 compares the distribution *pdfs* and quantiles of the random walk tangent part $t = \boldsymbol{\mu}^T \mathbf{x}_n$ with the asymptotic one for $n = 10$ steps on the $p = 3$ dimensional sphere. As explained in [15, p. 168–170], the *pdf* of the tangent part t can be derived from the symmetric directional *pdf* $f(\mathbf{x}) = g(\boldsymbol{\mu}^T \mathbf{x})$ on \mathbb{S}^{p-1} as

$$f_p(t) = \omega_{p-1} B\left(\frac{p-1}{2}, \frac{1}{2}\right)^{-1} g(t) (1-t^2)^{\frac{p-3}{2}},$$

for all $-1 \leq t \leq 1$, where $B(\cdot, \cdot)$ is the classical Beta function. In dimensions $p = 3$, the Fourier expansion (19) leads to the following simple expression for the projected distribution density, i.e. the tangent part *pdf*

$$f_3(t) = \sum_{\ell \geq 0} \frac{2\ell + 1}{2} \widehat{g}_\ell P_\ell(t),$$

for all $-1 \leq t \leq 1$, with \widehat{g}_ℓ the Fourier coefficient of the directional distribution *pdf* g . According to (20) and (36), $\widehat{g}_\ell = \widehat{f}_\ell^{\otimes n} = \left(\frac{I_{1/2+\ell}(\kappa)}{I_{1/2}(\kappa)} \right)^n$ for the n -step random walk on \mathbb{S}^2 . This Fourier expansion allows us to numerically evaluate the exact projected *pdf*. The empirical quantiles are estimated from 10^7 Monte-Carlo runs.

As expected, Fig. 1(a) shows that for low concentration, the asymptotic distribution diverges from the real one. However for high enough concentration, Figs 1(b) and especially 1(c) show the accuracy of the asymptotic approximation. In practice, when $p = 3$ the asymptotic distribution gives a reasonable approximation when $\tilde{\kappa}_n \approx \frac{\kappa}{n} \geq 10$ and is accurate when $\tilde{\kappa}_n \approx \frac{\kappa}{n} \geq 50$.

C. VMF multiple scattering process

The number of scattering events $N(t)$, which corresponds to the number of steps in the vMF random walk, is now assumed to be distributed as a Cox process with mixing variable $\Lambda(t)$: $N(t)|\Lambda(t) \sim \mathcal{P}(\Lambda_t)$.

1) *Asymptotic approximation of the vMF multiple scattering process*: Based on the vMF asymptotic distribution for the vMF random walk given in Thm. IV.1, it is possible to obtain an asymptotic distribution for the vMF scattering process

Proposition IV.2. Assume that there exists $a > 0$ such that $E[\Lambda(t)^a] < +\infty$. In the large κ case, an asymptotic expression for the *pdf* of the direction \mathbf{x}_t given in (30) for the vMF multiple scattering process is given by the following linear mixture

$$\tilde{f}(\mathbf{x}_t; \boldsymbol{\mu}, \kappa) = \mathcal{P}_0[f_{\Lambda_t}] \delta_{\boldsymbol{\mu}}(\mathbf{x}_t) + \sum_{n \geq 1} \mathcal{P}_n[f_{\Lambda_t}] f(\mathbf{x}_t; \boldsymbol{\mu}, \tilde{\kappa}_n), \quad (38)$$

where $f(\mathbf{x}_t; \boldsymbol{\mu}, \tilde{\kappa}_n)$ is the *pdf* of the vMF distribution $M_p(\boldsymbol{\mu}, \tilde{\kappa}_n)$ given in Thm. IV.1.

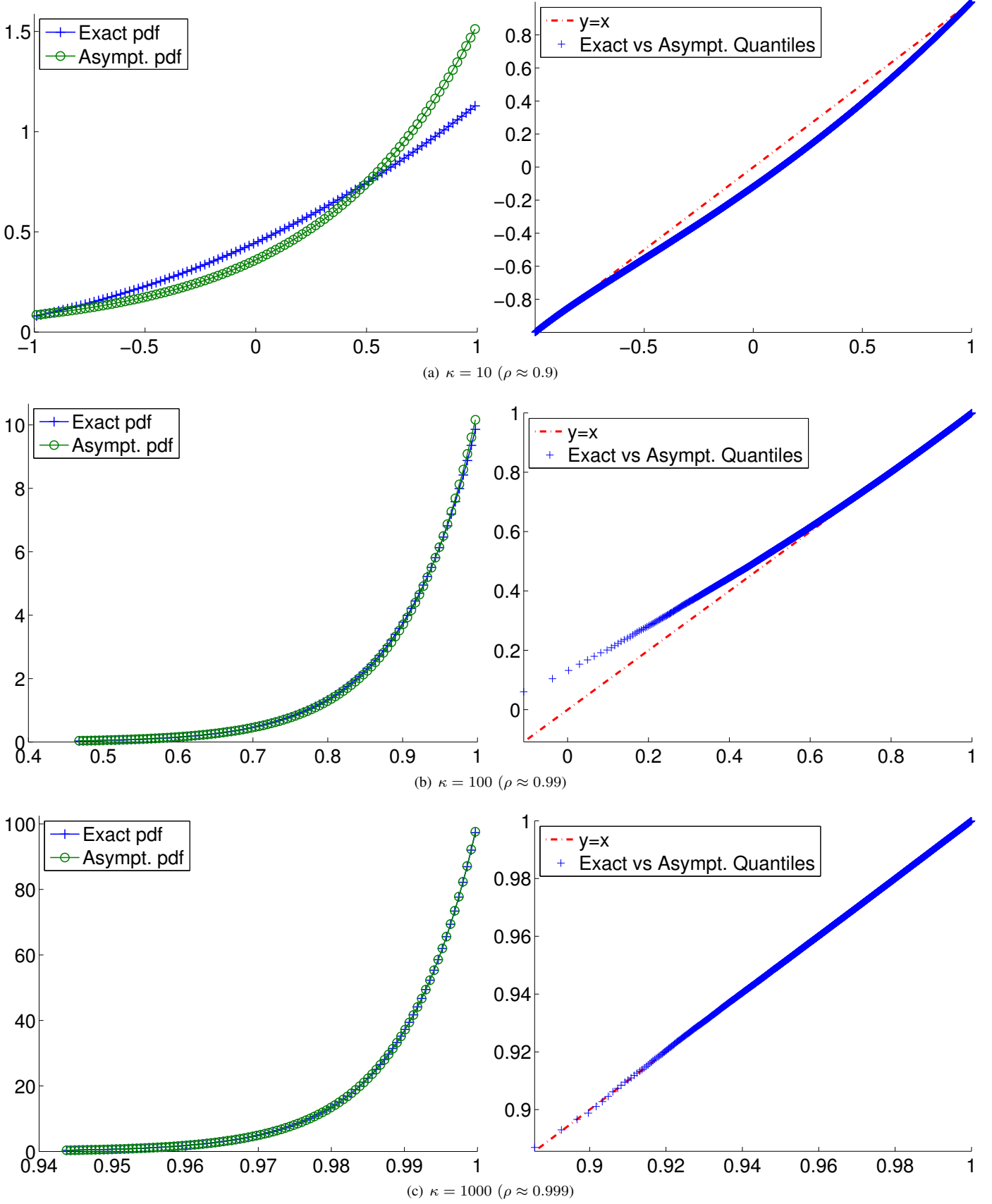


Fig. 1. Comparisons of the exact and asymptotic projected distributions for the $n = 10$ steps vMF random walk on \mathbb{S}^2 (i.e. $p = 3$) with concentration parameter κ . Top row: $\kappa = 10$, middle row: $\kappa = 100$, bottom row: $\kappa = 1000$. Left column: *pdfs* for the exact (blue cross) and asymptotic (green circles) distributions. Right column: empirical qq-plots for the exact distribution quantiles vs the asymptotic ones (10^7 samples).

Moreover, this yields the following approximation of the Fourier coefficients:

$$\widehat{h^{\otimes > 0}}_\ell = \widetilde{h}_\ell^{\otimes > 0} + O\left(\left(\frac{1}{\kappa}\right)^{\frac{3a}{a+3}}\right),$$

where $\widehat{h^{\otimes > 0}}_\ell$ is the ℓ th-order Fourier coefficient of the continuous part of the \mathbf{x}_t distribution given in (28), while $\widetilde{h}_\ell^{\otimes > 0}$ denotes the ℓ th-order Fourier coefficient of the continuous part of the asymptotic distribution.

Proof: The asymptotic pdf is obtained by plugging the vMF asymptotic distributions given in Thm. IV.1 in the multiple scattering mixture pdf given in (23). To show that the distributions are asymptotically equivalent, it is sufficient to show that their Fourier coefficients are asymptotically equivalent.

The expression of the ℓ th order Fourier coefficient of the continuous part of the density of \mathbf{x}_t given in (28) can be splitted in two terms:

$$\widehat{h^{\otimes > 0}}_\ell = \sum_{n=1}^{m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] \widehat{f}_\ell^{\otimes n} + \sum_{n>m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] \widehat{f}_\ell^{\otimes n}, \quad (39)$$

According to Thm IV.1, $\widehat{f}_\ell^{\otimes n} = \widetilde{f}_\ell^n + O\left(\left(\frac{m_\kappa}{\kappa}\right)^3\right)$, for all $n \leq m_\kappa$, thus the finite sum in the right-hand side of (39) can be dominated as

$$\begin{aligned} \sum_{n=1}^{m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] \widehat{f}_\ell^{\otimes n} &= \sum_{n=1}^{m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] \widetilde{f}_\ell^n + \sum_{n=1}^{m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] \times O\left(\left(\frac{m_\kappa}{\kappa}\right)^3\right), \\ &= \sum_{n=1}^{m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] \widetilde{f}_\ell^n + O\left(\left(\frac{m_\kappa}{\kappa}\right)^3\right), \end{aligned} \quad (40)$$

since $\sum_{n=1}^{m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] \leq \sum_{n \geq 1} \mathcal{P}_n[f_{\Lambda_t}] \leq 1$.

The second series in the right-hand side of (39) can be expressed as

$$\sum_{n>m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] \widehat{f}_\ell^{\otimes n} = \sum_{n>m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] \widetilde{f}_\ell^n + \sum_{n>m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] (\widehat{f}_\ell^{\otimes n} - \widetilde{f}_\ell^n), \quad (41)$$

with

$$\left| \sum_{n>m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] (\widehat{f}_\ell^{\otimes n} - \widetilde{f}_\ell^n) \right| \leq \sum_{n>m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] (|\widehat{f}_\ell^{\otimes n}| + |\widetilde{f}_\ell^n|) \leq 2 \sum_{n>m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] = 2 \Pr(N_t > m_\kappa),$$

where N_t is the Cox process counting the scattering events whose mixing intensity variable is Λ_t . A classical result about mixed Poisson distribution [27] yields that

$$\Pr(N_t > m_\kappa) = \int_0^{+\infty} e^{-\lambda_t} \frac{\lambda_t^{m_\kappa}}{m_\kappa!} \Pr(\Lambda_t > \lambda_t) d\lambda_t.$$

Because Λ_t is a positive random variable, the Markov inequality ensures that for $a > 0$

$$\Pr(\Lambda_t > \lambda_t) \leq \frac{E[\Lambda_t^a]}{\lambda_t^a}, \quad \text{for all } \lambda_t > 0.$$

As a consequence,

$$\Pr(N_t > m_\kappa) \leq E[\Lambda_t^a] \int_0^{+\infty} e^{-\lambda_t} \frac{\lambda_t^{m_\kappa - a}}{m_\kappa!} d\lambda_t = E[\Lambda_t^a] \frac{\Gamma(m_\kappa - a + 1)}{\Gamma(m_\kappa + 1)},$$

thus for large m_κ , $\Pr(N_t > m_\kappa) = O\left(\left(\frac{1}{m_\kappa}\right)^a\right)$, according to both the Stirling formula and the assumption that $E[\Lambda_t^a] < +\infty$. Eq. (41) can then be rewritten as

$$\sum_{n>m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] \widehat{f}_\ell^{\otimes n} = \sum_{n>m_\kappa} \mathcal{P}_n[f_{\Lambda_t}] \widetilde{f}_\ell^n + O\left(\left(\frac{1}{m_\kappa}\right)^a\right). \quad (42)$$

Plugging now (40) and (42) in (39) yields that

$$\widehat{h^{\otimes > 0}}_\ell = \widetilde{h}_\ell^{\otimes > 0} + O\left(\left(\frac{m_\kappa}{\kappa}\right)^3\right) + O\left(\left(\frac{1}{m_\kappa}\right)^a\right),$$

where $\widetilde{h}_\ell^{\otimes > 0} = \sum_{n \geq 1} \mathcal{P}_n[f_{\Lambda_t}] \widetilde{f}_\ell^n$ is the ℓ th order Fourier coefficient of $\sum_{n \geq 1} \mathcal{P}_n[f_{\Lambda_t}] f(\mathbf{x}_t; \boldsymbol{\mu}, \tilde{\kappa}_n)$ due to the orthogonality of the Legendre polynomials. Finally, setting $m_\kappa = \lfloor \kappa^\gamma \rfloor$ with $\gamma = \frac{3}{3+a} \in (0, 1)$ gives the expected result. This concludes the proof. \blacksquare

Note that when all the moments of the intensity variable Λ_t exist - this is the case for the Poisson process (Λ_t is deterministic) or the negative binomial process (Λ_t obeys a Gamma distribution) - the asymptotic approximation given in Prop. IV.2 yields an almost third order approximation of the Fourier coefficient for high concentration parameter κ .

From a practical point of view, it is also interesting to note that the asymptotic *pdf* approximation given in Prop. IV.2 may give a numerically simpler way to evaluate the *pdf* in the vMF case than the Fourier expansion (30). In fact, this mixture expression gives a series expansion with positive terms and weights. In addition, the mixture model offers a very simple way to draw some random variables asymptotically distributed according to the vMF multiple scattering process for large enough concentration parameter κ .

2) *Estimation bounds*: From the estimation theory perspective, it is of special interest to quantify the amount of information that the observed process carries about its distribution. In a parametric framework, the information about the parameters that govern the distribution is measured by the Fisher information matrix. Inverting the Fisher information provides now the Cramer-Rao lower bound (CRLB). This is a lower bound on the variance of any unbiased estimators of the parameters to be estimated. In addition, standard maximum likelihood estimators (MLEs) are known to be asymptotically unbiased and efficient under mild regularity conditions. This means that their large sample asymptotic variances approximately equal the CRLB.

The *pdf* expression of the multiple scattering process given in Prop. III.1, makes now possible to compute the Fisher information, and therefore the CRLB. However the Fisher information matrix requires to determine the covariance of the first-order derivatives with respect to the process parameters (or equivalently the negative expectations of second-order derivatives) of the log *pdf*. Closed-form expressions for both the Fisher information or the CRLB are difficult to obtain as this log *pdf* has no simple tractable expression. In such situation, it is very usual to approximate the expectations by using Monte Carlo methods. This allows one to numerically evaluate the CRLB based on the *pdf* Fourier expansion (30) for the vMF multiple scattering process. Moreover, in the high concentration case, i.e. for large κ , the vMF mixture representation given in Prop. IV.2 yields another simple way to approximate the CRLB with Monte Carlo methods.

The vMF concentration parameter $\kappa > 0$ and the mean resultant length $\rho \in (0, 1)$ of the vMF random steps are related by a one to one transformation $\rho = A_p(\kappa)$ given in (35). Therefore the random steps are reparametrized in the remainder by the scalar ρ . This yields a simple interpretation: the closer is ρ to 1, the more concentrated is the distribution, the closer to 0, the more uniform distribution. The CRLB for the parameter ρ reduces to $\frac{[A'_p(\kappa)]^2}{I(\kappa)}$ where $I(\kappa)$ is the Fisher information for κ .

When the number $N(t)$ of scattering events is a Poisson process, the intensity variable of the compound Cox process Λ_t is a deterministic value $\Lambda_t = \lambda_t$. The distribution of the multiple scattering process \mathbf{x}_t is then parametrized by the vector (ρ, λ_t) . Figs 2(a) and 2(b) depict the CRLBs for the parameters ρ and λ_t respectively as a function of $\kappa = A_p^{-1}(\rho)$ when the dimension is $p = 3$. The value of the Poisson intensity is set to $\lambda_t = 10$. One can see that the CRLBs quickly increase when the concentration decreases. In fact, in the limit case where ρ tends to zero (or equivalently, κ tends to zero), the distribution of \mathbf{x}_t converges to an uniform distribution on S^{p-1} and the model is not identifiable anymore. For larger values of ρ (i.e. for high κ), good estimation performances can be reached. Moreover there is little gain possible for the parameter λ_t that governs the number $N(t)$ of scattering events when ρ continues to converge to 1. Finally, one can see that the CRLBs computed for the high concentration asymptotic distribution of \mathbf{x}_t given in Prop. IV.2 are in good agreement with the exact one when ρ is close to 1.

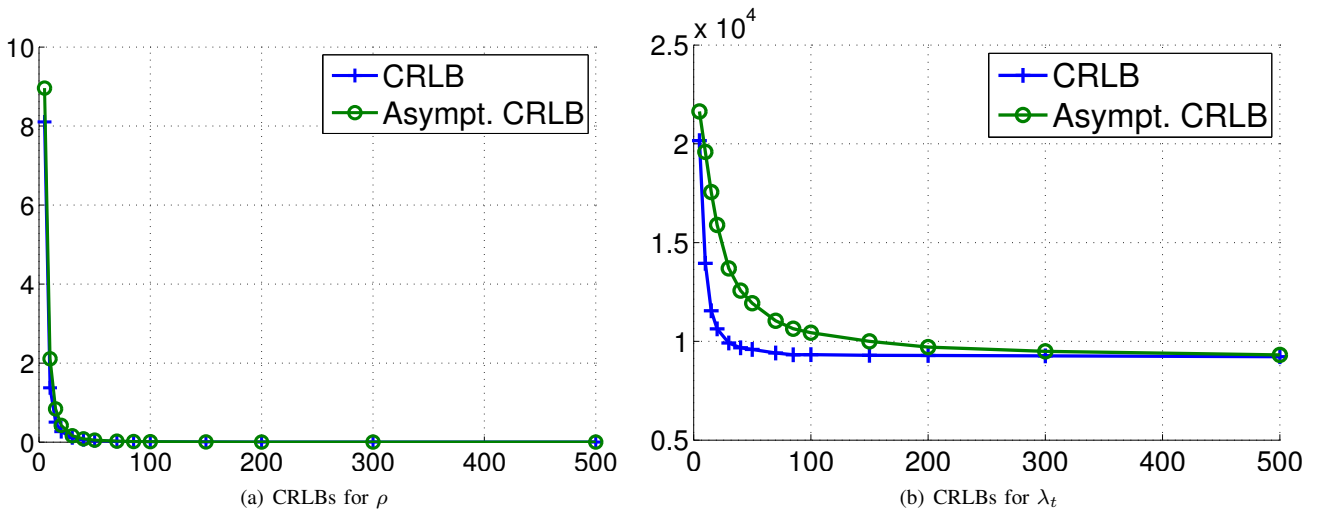


Fig. 2. CRLBs vs $\kappa = A_p^{-1}(\rho)$ for the Poisson multiple scattering process ($\lambda_t = 10$, $p = 3$). Blue curve: exact distribution of \mathbf{x}_t . Green curve: high concentration asymptotic distribution of \mathbf{x}_t .

Similar results are shown in Fig. 3 when the mean resultant length is set to a fixed value $\rho = 0.99$ (i.e. $\kappa \approx 100$) while

the intensity λ_t of the Poisson process varies. Fig. 3(a) shows that for too high λ_t , the model becomes hardly identifiable (in this case the distribution of \mathbf{x}_t converges to an uniform distribution). When ρ is fixed and for small enough λ_t , the high concentration asymptotic CRLB is in good agreement with the exact one.

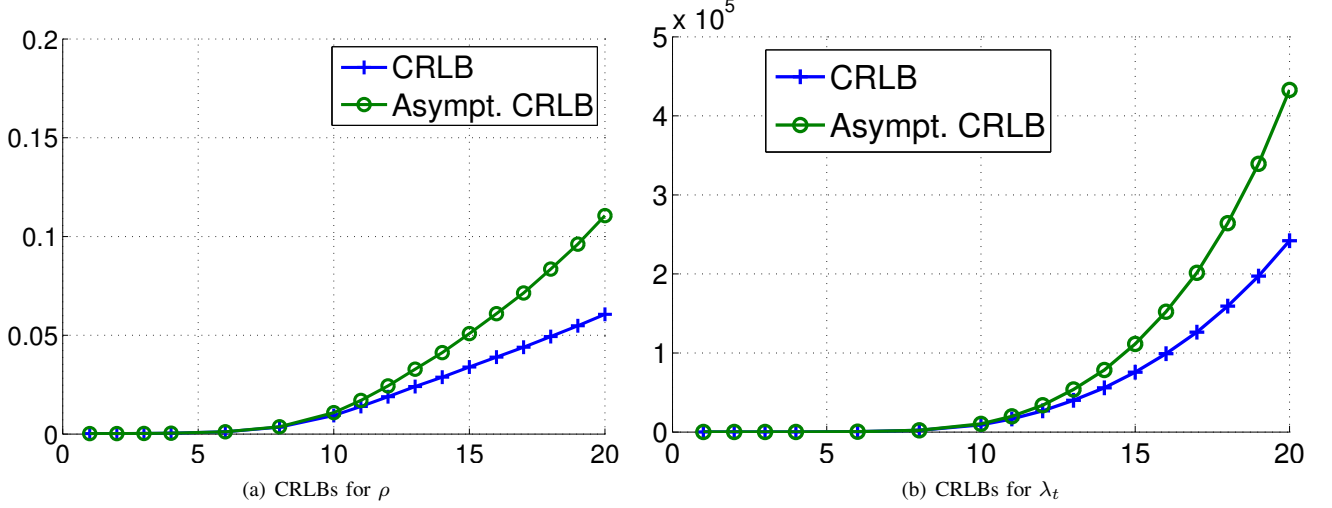


Fig. 3. CRLBs vs λ_t for the Poisson multiple scattering process ($\rho = 0.99$). Blue curve: exact distribution of \mathbf{x}_t . Green curve: high concentration asymptotic distribution of \mathbf{x}_t .

To conclude, we consider the more general case where the intensity is distributed as a Gamma process $\Lambda_t \sim \mathcal{G}(\xi_t, \theta)$, thus $N(t)$ is a Negative Binomial process parameterized by the vector (ρ, θ, ξ_t) . Figs 4(a), 4(b) and 4(c) depict the CRLBs for the parameters ρ , θ , and ξ_t respectively as a function of $\kappa = A_p^{-1}(\rho)$ when the dimension is $p = 3$. The parameters are set to $\theta = 1$ and $\xi_t = 10$. Similar conclusions to the one reported for Fig. 2 can be drawn.

The CRLBs presented for both Poisson and Negative Binomial cases outline the accuracy of the multiple scattering process model when the number of scattering events is low, which indicates its appropriateness and usefulness in forward scattering regimes occurring well before the full diffusion regime.

V. CONCLUSION

In this paper, we have studied multiple isotropic random walks and scattering processes on hyperspheres and obtained Fourier expansions for their *pdfs*. The case where the random steps follow a von Mises Fisher law on \mathbb{S}^{p-1} has been detailed and asymptotic approximation for multi-convolution of such densities have been introduced. The obtained expressions allow to numerically compute lower estimation bounds for the parameters of multiple scattering processes. These bounds should be of interest for future studies of estimation techniques (Method of moments, Bayesian, etc.) for such processes. The abundance of multiple scattering situations in engineering applications should provide applications for the presented results.

ACKNOWLEDGMENTS

The authors are grateful to Pr. Peter Jupp for fruitful discussions regarding directional statistics and convolution on double coset spaces.

APPENDIX

A. Proof of the unimodality of the convolution of unimodal and rotationally symmetric distributions

Let f_1 and f_2 be the *pdfs* of two absolutely continuous unimodal and rotationally symmetric distributions on \mathbb{S}^{p-1} with the same mode $\boldsymbol{\mu} \in \mathbb{S}^{p-1}$. These *pdfs* express as continuous functions of the only cosine $\mathbf{x}^T \boldsymbol{\mu}$, that is $f_1(\mathbf{x}; \boldsymbol{\mu}) = g_1(\mathbf{x}^T \boldsymbol{\mu})$ and $f_2(\mathbf{x}; \boldsymbol{\mu}) = g_2(\mathbf{x}^T \boldsymbol{\mu})$. Moreover, due to the unimodality property, g_1 and g_2 are increasing functions from $[-1, 1]$ to \mathbb{R}^+ . As explained in II-D, f_1 and f_2 belong to the double coset space $L^1(SO(p-1) \backslash SO(p) / SO(p-1), \mathbb{R})$ where $SO(p-1)$ stands for the rotation subgroup such that the axis defined by the unit vector $\boldsymbol{\mu}$ is left invariant. This space is stable by convolution, and the resulting convolved *pdf* is rotationally symmetric about $\boldsymbol{\mu}$ and reads

$$\begin{aligned} f(\mathbf{x}) &= (f_2 \star_{\boldsymbol{\mu}} f_1)(\mathbf{x}) = (f_2 \star f_1)(\mathbf{x}^T \boldsymbol{\mu}) \\ &= \int_{\mathbb{S}^{p-1}} g_2(\mathbf{x}^T \mathbf{x}_1) g_1(\mathbf{x}_1^T \boldsymbol{\mu}) d\mathbf{x}_1. \end{aligned}$$

We want to show now that the convolved distribution is unimodal with mode $\boldsymbol{\mu}$.

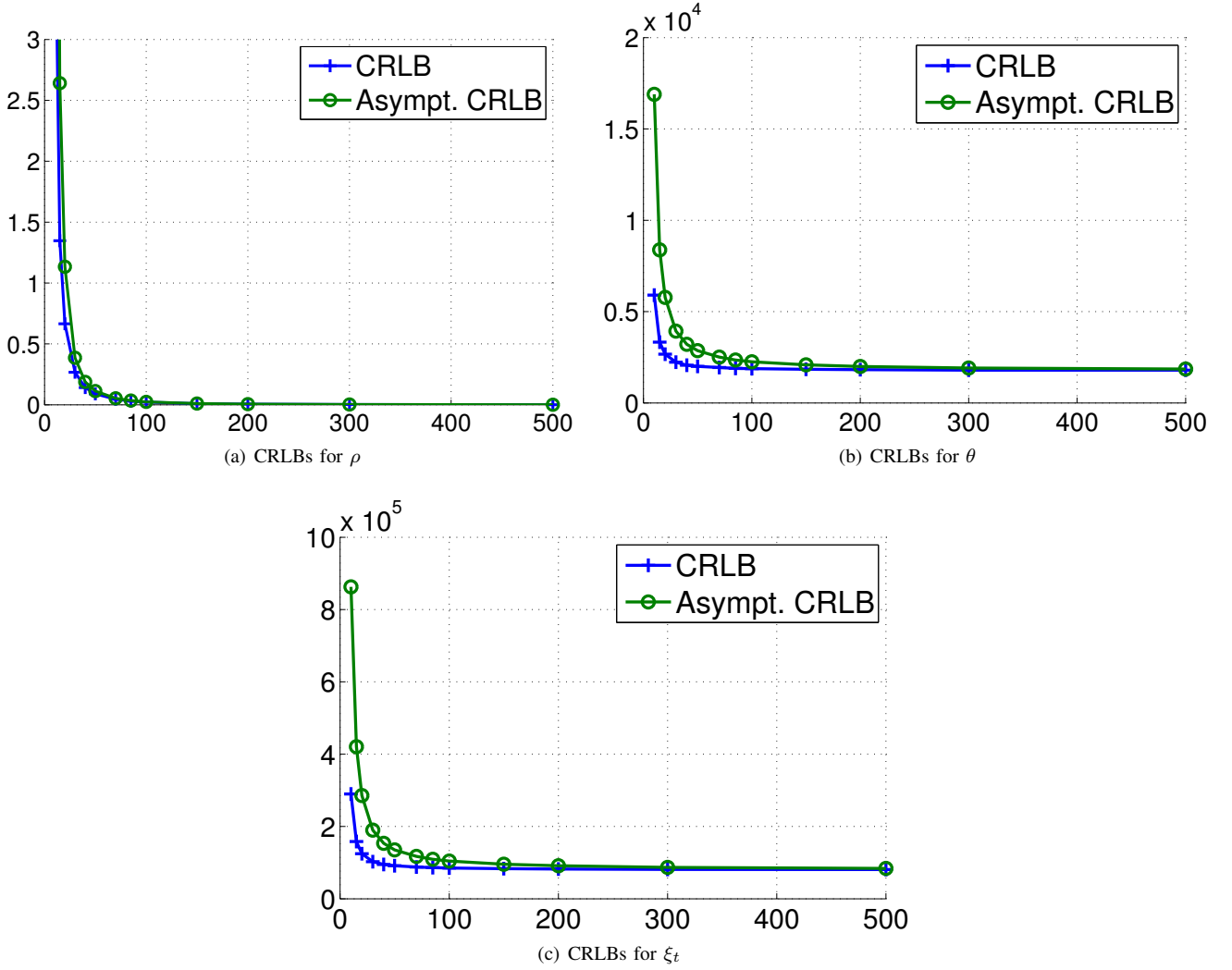


Fig. 4. CRLBs vs $\kappa = A_p^{-1}(\rho)$ for the Negative Binomial multiple scattering process ($\theta = 1$, $\xi_t = 10$, $p = 3$). Blue curve: exact distribution of \mathbf{x}_t . Green curve: high concentration asymptotic distribution of \mathbf{x}_t .

The sketch of the proof is inspired by the proof given in [28] of the equivalent property on the real line. Note first that the convolved *pdf* can be expressed as the following expectation:

$$\begin{aligned} f(\mathbf{x}) &= \int_{\mathbb{S}^{p-1}} g_2(\mathbf{x}^T \mathbf{x}_1) g_1(\mathbf{x}_1^T \boldsymbol{\mu}) d\mathbf{x}_1, \\ &= E \left[g_2(\mathbf{X}_1^T \mathbf{x}) \right], \end{aligned} \quad (43)$$

where \mathbf{X}_1 is a random vector in \mathbb{S}^{p-1} with *pdf* f_1 . Since $g_2(\mathbf{X}_1^T \mathbf{x})$ is a continuous positive variable, one gets

$$E \left[g_2(\mathbf{X}_1^T \mathbf{x}) \right] = \int_0^{+\infty} \Pr(g_2(\mathbf{X}_1^T \mathbf{x}) \geq u) du, \quad (44)$$

$$(45)$$

Furthermore, since g_2 is continuous and increasing on $(-1, 1)$, for all u in the image of g_2

$$\begin{aligned} \Pr \left(g_2(\mathbf{X}_1^T \mathbf{x}) \geq u \right) &= \Pr \left(\mathbf{X}_1^T \mathbf{x} \geq \delta \right), \\ &= \int_{\mathbf{x}_1 \in \mathbb{S}^{p-1} \cap \mathbf{x}^T \mathbf{x}_1 \geq \delta} g_1(\boldsymbol{\mu}^T \mathbf{x}_1) d\mathbf{x}_1. \end{aligned}$$

where $\delta = g_2^{-1}(u) \in (-1, 1)$. In the remainder, since we only consider unit vectors, the hypersphere constraint $\mathbf{x}_1 \in \mathbb{S}^{p-1}$ on the integration domain will be omitted to simplify the notation.

The normal-tangent decomposition along the axis $\boldsymbol{\mu}$ yields now that $\mathbf{x} = t\boldsymbol{\mu} + \sqrt{1-t^2}\boldsymbol{\xi}$ where $t \in [-1, 1]$ and $\boldsymbol{\xi} \in \boldsymbol{\mu}^\perp \cap \mathbb{S}^{p-1}$.

Lemma 2. For any fixed $\xi \in \mu^\perp \cap S^{p-1}$ and $\delta \in (-1, 1)$, the function h defined as

$$t \in [-1, 1] \mapsto h(t) = \int_{\mathbf{x}_t^T \mathbf{x}_1 \geq \delta} g_1(\mu^T \mathbf{x}_1) d\mathbf{x}_1,$$

where $\mathbf{x}_t = t\mu + \sqrt{1-t^2}\xi$ is increasing.

Proof: See section B. ■

According to Lemma 2, if $\mathbf{x}' = t'\mu + \sqrt{1-t'^2}\xi$ with $-1 \leq t < t' \leq 1$, then $\Pr(\mathbf{X}_1^T \mathbf{x} \geq \delta) < \Pr(\mathbf{X}_1^T \mathbf{x}' \geq \delta)$ for all $\delta \in (-1, 1)$. This yields directly that $f(\mathbf{x}) < f(\mathbf{x}')$ according to (45) and (43). Due to the rotational symmetry about μ , the convolved density expresses as $f(\mathbf{x}) = g(\mu^T \mathbf{x})$ and depends on the only tangent part $t = \mu^T \mathbf{x}$. Thus the inequality $f(\mathbf{x}) < f(\mathbf{x}')$ extends to the case where $\mathbf{x}' = t'\mu + \sqrt{1-t'^2}\xi'$ for any $\xi' \in \mu^\perp \cap S^{p-1}$. Finally, we obtain that g is an increasing function on $[-1, 1]$: for all $\mathbf{x}, \mathbf{x}' \in S^{p-1}$, $f(\mathbf{x}) = g(\mu^T \mathbf{x}) < f(\mathbf{x}') = g(\mu^T \mathbf{x}')$ iff $\mu^T \mathbf{x} < \mu^T \mathbf{x}'$, and the maximum is reached for $\mu^T \mathbf{x}' = 1$, i.e. when $\mathbf{x}' = \mu$. This concludes the proof.

B. Proof of Lemma 2

Consider the following unit vectors,

$$\begin{aligned} \mathbf{x}_t &= t\mu + \sqrt{1-t^2}\xi, \\ \mathbf{x}_{t'} &= t'\mu + \sqrt{1-t'^2}\xi, \end{aligned}$$

with $-1 \leq t < t' \leq 1$ and $\xi \in \mu^\perp \cap S^{p-1}$. Introduce the reflection matrix $R \in O(p)$ across the axis directed by $\mathbf{x}_t + \mathbf{x}_{t'}$, $R = \left[\frac{(\mathbf{x}_t + \mathbf{x}_{t'}) (\mathbf{x}_t + \mathbf{x}_{t'})^T}{1 + \mathbf{x}_t^T \mathbf{x}_{t'}} - I_p \right]$, where I_p is the $p \times p$ identity matrix.

This interchanges \mathbf{x}_t and $\mathbf{x}_{t'}$, i.e. $R\mathbf{x}_t = \mathbf{x}_{t'}$ and $R\mathbf{x}_{t'} = \mathbf{x}_t$. Let $\mathbf{u}_1 \in \mathbb{S}^{p-1}$ be defined as $\mathbf{u}_1 = R\mathbf{x}_1$. Since the reflection matrix R satisfies $R = R^T$, it comes that

$$\begin{aligned} \mathbf{x}_{t'}^T \mathbf{x}_1 &= (R\mathbf{x}_t)^T \mathbf{x}_1 = \mathbf{x}_t^T R\mathbf{x}_1 = \mathbf{x}_t^T \mathbf{u}_1, \\ \mathbf{x}_t^T \mathbf{x}_1 &= (R\mathbf{x}_{t'})^T \mathbf{x}_1 = \mathbf{x}_{t'}^T R\mathbf{x}_1 = \mathbf{x}_{t'}^T \mathbf{u}_1. \end{aligned} \tag{46}$$

Moreover one gets that

$$h(t) = \int_{\substack{\mathbf{x}_t^T \mathbf{x}_1 \geq \delta \\ \cap \mathbf{x}_{t'}^T \mathbf{x}_1 \geq \delta}} g_1(\mu^T \mathbf{x}_1) d\mathbf{x}_1 + \underbrace{\int_{\substack{\mathbf{x}_t^T \mathbf{x}_1 \geq \delta \\ \cap \mathbf{x}_{t'}^T \mathbf{x}_1 < \delta}} g_1(\mu^T \mathbf{x}_1) d\mathbf{x}_1}_I.$$

The reflection matrix R satisfies $R\mathbf{x}_t = \mathbf{x}_{t'}$, $|\det R| = 1$, $R^{-1} = R$ and (46). As a consequence, performing the substitution $\mathbf{u}_1 = R\mathbf{x}_1$ in the integral denoted as I yields

$$I = \int_{\substack{\mathbf{x}_{t'}^T \mathbf{u}_1 \geq \delta \\ \cap \mathbf{x}_t^T \mathbf{u}_1 < \delta}} g_1(\mu^T R\mathbf{u}_1) d\mathbf{u}_1. \tag{47}$$

We need now to use the following result

Lemma 3. For all $\mathbf{x}_1 \in \mathbb{S}^{p-1}$ such that $\mathbf{x}_t^T \mathbf{x}_1 > \mathbf{x}_{t'}^T \mathbf{x}_1$,

$$\mu^T \mathbf{u}_1 > \mu^T \mathbf{x}_1. \tag{48}$$

Proof of Lemma 3: When $\mathbf{x}_t^T \mathbf{x}_1 > \mathbf{x}_{t'}^T \mathbf{x}_1$, eq. (46) allows us to derive the following inequalities

$$\begin{aligned} \mathbf{x}_{t'}^T \mathbf{u}_1 &> \mathbf{x}_{t'}^T \mathbf{x}_1, \\ \mathbf{x}_t^T \mathbf{x}_1 &> \mathbf{x}_t^T \mathbf{u}_1. \end{aligned}$$

Using the normal tangent decomposition, theses inequalities express as

$$t'\mu^T \mathbf{u}_1 + \sqrt{1-t'^2}\xi^T \mathbf{u}_1 > t'\mu^T \mathbf{x}_1 + \sqrt{1-t'^2}\xi^T \mathbf{x}_1, \tag{49}$$

$$t\mu^T \mathbf{x}_1 + \sqrt{1-t^2}\xi^T \mathbf{x}_1 > t\mu^T \mathbf{u}_1 + \sqrt{1-t^2}\xi^T \mathbf{u}_1. \tag{50}$$

When $t' = 1$, or $t = -1$ respectively, it comes directly from (49), or (50) respectively, that $\mu^T \mathbf{u}_1 > \mu^T \mathbf{x}_1$. We can thus assume that $-1 < t < t' < 1$. Multiplying both sides of inequality (50) and (49) by $\frac{1}{\sqrt{1-t^2}} > 0$ and $\frac{-1}{\sqrt{1-t'^2}} < 0$ respectively, and summing the resulting inequalities yields

$$(\alpha(t') - \alpha(t))\mu^T \mathbf{u}_1 > (\alpha(t') - \alpha(t))\mu^T \mathbf{x}_1,$$

where $\alpha(z) = \frac{z}{\sqrt{1-z^2}}$. The function α being increasing on $(-1, 1)$, the factor $\alpha(t') - \alpha(t)$ is positive since $-1 < t < t' < 1$. Thus the required inequality holds. ■

For all \mathbf{u}_1 belonging to the integration domain defined in (47), it comes that $\mathbf{x}_{t'}^T \mathbf{u}_1 > \mathbf{x}_t^T \mathbf{u}_1$ which is equivalent to $\mathbf{x}_t^T \mathbf{x}_1 > \mathbf{x}_{t'}^T \mathbf{x}_1$. Thus $\boldsymbol{\mu}^T R \mathbf{u}_1 = \boldsymbol{\mu}^T \mathbf{x}_1 < \boldsymbol{\mu}^T \mathbf{u}_1$ according to Lemma 3. Since g_1 is increasing, it comes that

$$I < \int_{\substack{\mathbf{x}_{t'}^T \mathbf{u}_1 \geq \delta \\ \cap \mathbf{x}_t^T \mathbf{u}_1 < \delta}} g_1(\boldsymbol{\mu}^T \mathbf{u}_1) d\mathbf{u}_1.$$

As a consequence,

$$\begin{aligned} h(t) &< \int_{\substack{\mathbf{x}_t^T \mathbf{x}_1 \geq \delta \\ \cap \mathbf{x}_{t'}^T \mathbf{x}_1 \geq \delta}} g_1(\boldsymbol{\mu}^T \mathbf{x}_1) d\mathbf{x}_1 + \int_{\substack{\mathbf{x}_{t'}^T \mathbf{u}_1 \geq \delta \\ \cap \mathbf{x}_t^T \mathbf{u}_1 < \delta}} g_1(\boldsymbol{\mu}^T \mathbf{u}_1) d\mathbf{u}_1, \\ &= \int_{\substack{\mathbf{x}_{t'}^T \mathbf{z} \geq \delta \\ \cap \mathbf{x}_t^T \mathbf{z} \geq \delta}} g_1(\boldsymbol{\mu}^T \mathbf{z}) d\mathbf{z} + \int_{\substack{\mathbf{x}_{t'}^T \mathbf{z} \geq \delta \\ \cap \mathbf{x}_t^T \mathbf{z} < \delta}} g_1(\boldsymbol{\mu}^T \mathbf{z}) d\mathbf{z}, \\ &= \int_{\mathbf{x}_{t'}^T \mathbf{z} \geq \delta} g_1(\boldsymbol{\mu}^T \mathbf{z}) d\mathbf{z} = h(t'), \end{aligned}$$

and the inequality holds for all $-1 \leq t < t' \leq 1$.

REFERENCES

- [1] A. Bhalerao and C.-. Westin, "Hyperspherical von mises-fisher mixture (hvmf) modelling of high angular resolution diffusion mri," in *MICCAI 2007*, N. Ayache, S. Ourselin, and A. Maeder, Eds., vol. 4791 of *Lecture Notes in Computer Science*, pp. 236–243. Springer Berlin Heidelberg, 2007.
- [2] M. Bangert, P. Hennig, and U. Oelfke, "Using an infinite von mises-fisher mixture to cluster treatment beam directions in external radiation therapy," in *Proceedings of the Ninth international conference on Machine Learning and applications*, 2010, pp. 746–751.
- [3] H. Tang and S.M. Chu, "Generative model-based speaker clustering via mixture of von mises-fisher distributions," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2009, pp. 4101–4104.
- [4] A. Banerjee, I.S. Dhillon, J. Ghosh, and S. Sra, "Clustering on the unit hypersphere using von mises-fisher distributions," *Journal of Machine Learning Research*, vol. 6, pp. 1345 – 1382, 2005.
- [5] X. Ning, L. Papiez, and G. Sandison, "Compound-poisson-process method for the multiple scattering of charged particles," *Phys. Rev. E*, vol. 52, no. 5, pp. 5621–5633, Nov 1995.
- [6] N. Le Bihan and L. Margerin, "Nonparametric estimation of the heterogeneity of a random medium using compound poisson process modeling of wave multiple scattering," *Physical Review E*, vol. 80, pp. 016601, 2009.
- [7] S. Said, C. Lageman, N. Le Bihan, and J.H. Manton, "Decompounding on compact Lie groups," *IEEE Transactions on Information theory*, vol. 56, no. 6, pp. 2766 – 2777, 2010.
- [8] F. Perrin, *Étude mathématique du mouvement Brownien de rotation*, Ph.D. thesis, Faculté des sciences de Paris, 1925.
- [9] I. Dokmanić and D. Petrinović, "Convolution on the n -sphere with application to pdf modeling," *Signal Processing, IEEE Transactions on*, vol. 58, no. 3, pp. 1157–1170, March 2010.
- [10] M. Franceschetti, J. Bruck, and L.J. Schulman, "A random walk model of wave propagation," *Antennas and Propagation, IEEE Transactions on*, vol. 52, no. 5, pp. 1304–1317, May 2004.
- [11] S. Durant, O. Calvo-Perez, N. Vukadinovic, and J.-J. Greffet, "Light scattering by a random distribution of particles embedded in absorbing media: full-wave Monte Carlo solutions of the extinction coefficient," *J. Opt. Soc. Am. A*, vol. 24, no. 9, pp. 2953–2962, September 2007.
- [12] C. Jin, R.R. Nadakuditi, E. Michielssen, and S. Rand, "An iterative, backscatter-analysis based algorithm for increasing transmission through a highly-backscattering random medium," in *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, Aug 2012, pp. 97–100.
- [13] M. Ghogho, O. Besson, and A. Swami, "Estimation of directions of arrival of multiple scattered sources," *Signal Processing, IEEE Transactions on*, vol. 49, no. 11, pp. 2467–2480, Nov 2001.
- [14] M. Costa, V. Koivunen, and H.V. Poor, "Estimating directional statistics using wavefield modeling and mixtures of von-mises distributions," *Signal Processing Letters, IEEE*, vol. 21, no. 12, pp. 1496–1500, Dec 2014.
- [15] K.V. Mardia and P.E. Jupp, *Directional statistics*, John Wiley & Sons Ltd, 2000.
- [16] J. Kent, "Limiting behaviour of the von Mises-Fisher distribution," *Math. Proc. Camb. Phil. Soc.*, vol. 84, pp. 531–536, 1978.
- [17] M. Volker, *Lectures on constructive approximation. Fourier, Spline, and wavelet methods on the real line, the sphere and the ball*, Birkhauser, 2013.
- [18] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, 1972.
- [19] H. Dym and H.P. McKean, *Fourier series and integrals*, Academic Press, 1972.
- [20] A. Ishimaru, *Wave propagation and scattering in random media*, Wiley-IEEE Press, 1999.
- [21] F. Chatelain and N. Le Bihan, "von-mises fisher approximation of multiple scattering process on the hypershpere," in *Icassp*, 2013.
- [22] M. Lefebvre, *Applied stochastic processes*, Springer, 2006.
- [23] B. Saleh, *Photoelectron statistics*, Springer, 1978.
- [24] A. Ferrari, G. Letac, and J.-Y. Tournet, "Exponential families of mixed Poisson distributions," *Journal of Multivariate Analysis*, vol. 98, no. 6, pp. 1283–1292, 2007.
- [25] F. Chatelain, S. Lambert-Lacroix, and J.-Y. Tournet, "Pairwise likelihood estimation for multivariate mixed poisson models generated by gamma intensities," *Statistics and Computing*, vol. 19, no. 3, pp. 283–301, Sept. 2009.
- [26] Patrick Billingsley, *Probability and Measure*, Wiley, New York, NY, 3rd edition, 1995.
- [27] Dimitris Karlis and Evdokia Xekalaki, "Mixed Poisson distributions," *International Statistical Review*, vol. 73, no. 1, pp. 35–58, 2005.
- [28] S. Purkayastha, "Simple proofs of two results on convolutions of unimodal distributions," *Statistics & Probability Letters*, vol. 39, no. 2, pp. 97–100, 1998.

Nicolas Le Bihan Nicolas Le Bihan obtained his B.Sc. degree in physics from the Université de Bretagne Occidentale in Brest, France, in 1997. He received the M.Sc. and Ph.D. degrees in signal processing in 1998 and 2001, respectively, both from Grenoble INP. In 2011, he obtained the Habilitation degree from Grenoble INP. Since 2002, he has been a research associate at the Centre National de la Recherche Scientifique (CNRS) and is working at the Department of Images and Signals of the GIPSA-Lab (CNRS UMR 5083) in Grenoble, France. From 2013 to 2015, he was a visiting fellow at the University of Melbourne, holding a Marie Curie International Outgoing Fellowship from the European Union (IOF GeoSToSip 326176), ERA, 7th PCRD. His research interests include statistical signal processing on groups, noncommutative algebras and differentiable manifolds and its applications in polarized wave physics, waves in disordered media, and geophysics.

Florent Chatelain Florent Chatelain received the Eng. degree in computer sciences and applied mathematics from ENSIMAG, Grenoble, France, and the M.Sc. degree in applied mathematics from the University Joseph Fourier of Grenoble, France, both in June 2004, and the Ph.D. degree in signal processing from the National Polytechnic Institute, Toulouse, France, in 2007. He is currently an Assistant Professor at GIPSA-Lab, University of Grenoble, France. His research interests are centered around estimation, detection, and the analysis of stochastic processes.

Jonathan Manton Professor Jonathan Manton holds a Distinguished Chair at the University of Melbourne with the title Future Generation Professor. He is also an Adjunct Professor in the Mathematical Sciences Institute at the Australian National University. Prof Manton is a Fellow of IEEE and a Fellow of the Australian Mathematical Society. He received his Bachelor of Science (mathematics) and Bachelor of Engineering (electrical) degrees in 1995 and his Ph.D. degree in 1998, all from the University of Melbourne, Australia. From 1998 to 2004, he was with the Department of Electrical and Electronic Engineering at the University of Melbourne. During that time, he held a Postdoctoral Research Fellowship then subsequently a Queen Elizabeth II Fellowship, both from the Australian Research Council. In 2005 he became a full Professor in the Department of Information Engineering, Research School of Information Sciences and Engineering (RSISE) at the Australian National University. From July 2006 till May 2008, he was on secondment to the Australian Research Council as Executive Director, Mathematics, Information and Communication Sciences. Prof Manton's traditional research interests range from pure mathematics (e.g. commutative algebra, algebraic geometry, differential geometry) to engineering (e.g. signal processing, wireless communications, systems theory). More recently, he has become interested in systems biology and systems neuroscience.